



10

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C07H 17/08, A61K 31/70</b>	<b>A2</b>	<b>(11) International Publication Number:</b> <b>WO 00/00500</b> <b>(43) International Publication Date:</b> 6 January 2000 (06.01.00)
<b>(21) International Application Number:</b> PCT/GB99/02042 <b>(22) International Filing Date:</b> 29 June 1999 (29.06.99)  <b>(30) Priority Data:</b> 9814006.4                      29 June 1998 (29.06.98)                      GB  <b>(71) Applicants (for all designated States except US):</b> BIOTICA TECHNOLOGY LIMITED [GB/GB]; 112 Hills Road, Cambridge CB2 1PH (GB). PFIZER INC. [US/US]; 235 East 42nd Street, New York, NY 10017 (US).  <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> LEADLAY, Peter, Francis [GB/GB]; 17 Clarendon Road, Cambridge CB2 2BH (GB). STAUNTON, James [GB/GB]; 29 Porson Road, Cambridge CB2 2ET (GB). CORTES, Jesus [GB/GB]; 26 Cambanks, Union Lane, Cambridge CB4 1PZ (GB). MCARTHUR, Hamish, Alastair, Irvine [GB/US]; 19 Pheasant Run Drive, Gales Ferry, CT 06335 (US).  <b>(74) Agents:</b> STUART, Ian et al.; Mewburn Ellis, York House, 23 Kingsway, London WC2B 6HP (GB).		<b>(81) Designated States:</b> AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i>
<b>(54) Title:</b> POLYKETIDES AND THEIR SYNTHESIS  <b>(57) Abstract</b> <p>A polyketide synthase ("PKS") of Type I is a complex multienzyme including a loading domain linked to a multiplicity of extension domains. This first extension module receives an acyl starter unit from the loading domain and each extension module adds a further ketide unit which may undergo processing (e.g. reduction). We have found that the Ksq domain possessed by some PKS's has decarboxylating activity, e.g. generating (substituted) acyl from (substituted) malonyl. The CLF domain of type II PKS's has similar activity. By inserting loading modules including such domains into PKS's not normally possessing them it is possible to control the starter units used.</p>		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LJ	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Polyketides and their Synthesis

The present invention relates to processes and materials (including enzyme systems, nucleic acids, vectors and cultures) for preparing 14-membered macrolides by recombinant synthesis and to the novel polyketides so produced. Polyketide biosynthetic genes or portions of them, which may be derived from different polyketide biosynthetic gene clusters are manipulated to allow the production of specific novel polyketides, such as 12-, 14- and 16-membered macrolides, of predicted structure. This invention is particularly concerned with the replacement of genetic material encoding the natural starter unit with other genes in order to prepare 14-membered macrolides with preferentially an acetate starter unit, whilst minimising the formation of by-products containing a different starter unit.

Polyketides are a large and structurally diverse class of natural products that includes many compounds possessing antibiotic or other pharmacological properties, such as erythromycin, tetracyclines, rapamycin, avermectin, monensin, epothilones and FK506. In particular, polyketides are abundantly produced by *Streptomyces* and related actinomycete bacteria. They are synthesised by the repeated stepwise condensation of acylthioesters in a manner analogous to that of fatty acid biosynthesis. The greater structural diversity found among natural polyketides arises from the selection of (usually) acetate or propionate as starter or "extender" units; and from the

differing degree of processing of the  $\beta$ -keto group observed after each condensation. Examples of processing steps include reduction to  $\beta$ -hydroxyacyl-, reduction followed by dehydration to 2-enoyl-, and complete reduction to the saturated acylthioester. The stereochemical outcome of these processing steps is also specified for each cycle of chain extension.

The biosynthesis of polyketides is initiated by a group of chain-forming enzymes known as polyketide synthases. Two classes of polyketide synthase (PKS) have been described in actinomycetes. One class, named Type I PKSs, represented by the PKSs for the macrolides erythromycin, oleandomycin, avermectin and rapamycin, consists of a different set or "module" of enzymes for each cycle of polyketide chain extension. For an example see Figure 1 (Cortés, J. et al. *Nature* (1990) 348:176-178; Donadio, S. et al. *Science* (1991) 2523:675-679; Swan, D.G. et al. *Mol. Gen. Genet.* (1994) 242:358-362; MacNeil, D. J. et al. *Gene* (1992) 115:119-125; Schwecke, T. et al. *Proc. Natl. Acad. Sci. USA* (1995) 92:7839-7843).

The term "extension module" as used herein refers to the set of contiguous domains, from a  $\beta$ -ketoacyl-ACP synthase ("AKS") domain to the next acyl carrier protein ("ACP") domain, which accomplishes one cycle of polyketide chain extension. The term "loading module" is used to refer to any group of contiguous domains which accomplishes the loading of the starter unit onto the PKS and thus renders it available to the KS domain of the first

extension module. The length of polyketide formed has been altered, in the case of erythromycin biosynthesis, by specific relocation using genetic engineering of the enzymatic domain of the erythromycin-producing PKS that contains the chain releasing thioesterase/cyclase activity (Cortés et al. Science (1995) 268:1487-1489; Kao, C.M. et al. J. Am. Chem. Soc. (1995) 117:9105-9106).

In-frame deletion of the DNA encoding part of the ketoreductase domain in module 5 of the erythromycin-producing PKS (also known as 6-deoxyerythronolide B synthase, DEBS) has been shown to lead to the formation of erythromycin analogues 5,6-dideoxy-3- $\alpha$ -mycarosyl-5-oxoerythronolide B, 5,6-dideoxy-5-oxoerythronolide B and 5,6-dideoxy, 6  $\beta$ -epoxy-5-oxoerythronolide B (Donadio, S. et al. Science (1991) 252:675-679). Likewise, alteration of active site residues in the enoylreductase domain of module 4 in DEBS, by genetic engineering of the corresponding PKS-encoding DNA and its introduction into *Saccharopolyspora erythraea*, led to the production of 6,7-anhydroerythromycin C (Donadio, S. et al. Proc Natl. Acad. Sci. USA (1993) 90:7119-7123).

International Patent Application number WO 93/13663 describes additional types of genetic manipulation of the DEBS genes that are capable of producing altered polyketides. However many such attempts are reported to have been unproductive (Hutchinson, C. R. and Fujii, I. Annu. Rev. Microbiol. (1995) 49:201-238, at p. 231). The complete DNA sequence of the genes from *Streptomyces*

*hygroscopicus* that encode the modular Type I PKS governing the biosynthesis of the macrocyclic immunosuppressant polyketide rapamycin has been disclosed (Schwecke, T. et al. (1995) Proc. Natl. Acad. Sci. USA 92:7839-7843). The DNA sequence is deposited in the EMBL/Genbank Database under the accession number X86780.

The second class of PKS, named Type II PKSs, is represented by the synthases for aromatic compounds. Type II PKSs contain only a single set of enzymatic activities for chain extension and these are re-used as appropriate in successive cycles (Bibb, M. J. et al. EMBO J. (1989) 8:2727-2736; Sherman, D. H. et al. EMBO J. (1989) 8:2717-2725; Fernandez-Moreno, M.A. et al. J. Biol. Chem. (1992) 267:19278-19290). The "extender" units for the Type II pKSs are usually acetate units, and the presence of specific cyclases dictates the preferred pathway for cyclisation of the completed chain into an aromatic product (Hutchinson, C. R. and Fujii, I. Annu. Rev. Microbiol. (1995) 49:201-238). Hybrid polyketides have been obtained by the introduction of clones Type II PKS gene-containing DNA into another strain containing a different Type II PKS gene cluster, for example by introduction of DNA derived from the gene cluster for actinorhodin, a blue-pigmented polyketide from *Streptomyces coelicolor*, into an anthraquinone polyketide-producing strain of *Streptomyces galileus* (Bartel, P. L. et al. J. Bacteriol. (1990) 172:4816-4826).

The minimal number of domains required for polyketide

chain extension on a Type II PKS when expressed in a *Streptomyces coelicolor* host cell (the "minimal PKS") has been defined for example in International Patent Application Number WO 95/08548 as containing the following

5 three polypeptides which are products of the *act I* genes: first KS; secondly a polypeptide termed the CLF with end-to-end amino acid sequence similarity to the KS but in which the essential active site residue of the KS, namely a cysteine residue, is substituted either by a glutamine

10 residue, or in the case of the PKS for a spore pigment such as the *whiE* gene product (Chater, K. F. and Davis, N. K. Mol. Microbiol. (1990) 4:1679-1691) by a glutamic acid residue (Figure 2); and finally an ACP. The CLF has been stated for example in International Patent Application

15 Number WO 95/08548 to be a factor that determines the chain length of the polyketide chain that is produced by the minimal PKS. However it has been found (Shen, B. et al. J. Am. Chem. Soc. (1995) 117:6811-6821) that when the CLF for the octaketide actinorhodin is used to replace the CLF for

20 the decaketide tetracenomycin in host cells of *Streptomyces glaucescens*, the polyketide product is not found to be altered from a decaketide to an octaketide, so the exact role of the CLF remains unclear. An alternative nomenclature has been proposed in which KS is designated

25 KS $\alpha$  and CLF is designated KS $\beta$ , to reflect this lack of knowledge (Meurer, G. et al. Chemistry and Biology (1997) 4:433-443). The mechanism by which acetate starter units and acetate extender units are loaded onto the Type II PKS

is not known, but it is speculated that the malonyl-CoA: ACP acyltransferase of the fatty acid synthase of the host cell can fulfil the same function for the Type II PKS (Revill, W. P. et al. J. Bacteriol. (1995) 177:3946-3952).

- 5 International Patent Application Number WO 95/08548 describes the replacement of actinorhodin PKS genes by heterologous DNA from other Type II PKS gene clusters, to obtain hybrid polyketides. The same International Patent Application WO 95/08548 describes the construction of a
- 10 strain of *Streptomyces coelicolor* which substantially lacks the native gene cluster for actinorhodin, and the use in that strain of a plasmid vector pRM5 derived from the low-copy number vector SCP2\* isolated from *Streptomyces coelicolor* (Bibb, M. J. and Hopwood, D. A. J. Gen.
- 15 Microbiol. (1981) 126:427) and in which heterologous PKS-encoding DNA may be expressed under the control of the divergent *act I*/*act III* promoter region of the actinorhodin gene cluster (Fernandez-Moreno, M.A. et al. J. Biol. Chem. (1992) 267:19278-19290). The plasmid pRM5 also
- 20 contains DNA from the actinorhodin biosynthetic gene cluster encoding the gene for a specific activator protein, ActII-orf4. The Act II-orf4 protein is required for transcription of the genes placed under the control of the *actI*/*act II* bidirectional promoter and activates gene
- 25 expression during the transition from growth to stationary phase in the vegetative mycelium (Hallam, S. E. et al. Gene (1988) 74:305-320).

Type II clusters in *Streptomyces* are known to be



activated by pathway-specific activator genes (Narva, K. E. and Feitelson, J. S. J. Bacteriol. (1990) 172:326-333; Stutzman-Engwall, K. J. et al. J. Bacteriol. (1992) 174:144-154; Fernandez-Moreno, M.A. et al. Cell (1991) 5 66:769-780; Takano, E. et al. Mol. Microbiol. (1992) 6:2797-2804; Takano, E. et al. Mol. Microbiol. (1992) 7:837-845), The DnrI gene product complements a mutation in the actII-orf4 gene of *S. coelicolor*, implying that DnrI and ActII-orf4 proteins act on similar targets. A gene 10 (*srmR*) has been described (EP 0 524 832 A2) that is located near the Type I PKS gene cluster for the macrolide polyketide spiramycin. This gene specifically activates the production of the macrolide antibiotic spiramycin, but no other examples have been found of such a gene. Also, no 15 homologues of the ActII-orf4/DnrI/RedD family of activators have been described that act on Type I PKS genes.

Although large numbers of therapeutically important polyketides have been identified, there remains a need to obtain novel polyketides that have enhanced properties or 20 possess completely novel bioactivity. The complex polyketides produced by Type I PKSs are particularly valuable, in that they include compounds with known utility as anthelmintics, insecticides, immunosuppressants, antifungal or antibacterial agents. Because of their 25 structural complexity, such novel polyketides are not readily obtainable by total chemical synthesis, or by chemical modifications of known polyketides.

There is also a need to develop reliable and specific ways of deploying individual modules in practice so that all, or a large fraction, of hybrid PKS genes that are constructed, are viable and produce the desired polyketide product.

Pending International Patent Application number PCT/GB97/01819 discloses that a PKS gene assembly (particularly of Type I) encodes a loading module which is followed by at least one extension module. Thus Figure 1 shows the organisation of the DEBS genes. The first open reading frame encodes the first multi-enzyme or cassette (DEBS 1) which consists of three modules: the loading module (ery-load) and two extension modules (modules 1 and 2). The loading module comprises an acyltransferase and an acyl carrier protein. This may be contrasted with Fig. 1 of WO 93/13663 (referred to above). This shows ORF1 to consist of only two modules, the first of which is in fact both the loading module and the first extension module.

PCT/GB97/01819 describes in general terms the production of a hybrid PKS gene assembly comprising a loading module and at least one extension module. PCT/GB97/01819 also describes (see also Marsden, A. F. A. et al. Science (1998) 279:199-202) construction of a hybrid PKS gene assembly by grafting the wide-specificity loading module for the avermectin-producing polyketide synthase onto the first multienzyme component (DEBS 1) for the erythromycin PKS in place of the normal loading module. Certain novel polyketides can be prepared using the hybrid PKS gene assembly, as described for example in pending

International Patent Application number (PCT/GB97/01810).  
Patent Application PCT/GB97/01819 further describes the  
construction of a hybrid PKS gene assembly by grafting the  
loading module for the rapamycin-producing polyketide  
5 synthase onto the first multienzyme component (DEBS 1) for  
the erythromycin PKS in place of the normal loading module.  
The loading module of the rapamycin PKS differs from the  
loading modules of DEBS and the avermectin PKS in that it  
comprises a CoA ligase domain, an enoylreductase ("ER")  
10 domain and an ACP, so that suitable organic acids including  
the natural starter unit 3,4-dihydroxycyclohexane  
carboxylic acid may be activated in situ on the PKS loading  
domain, and with or without reduction by the ER domain  
transferred to the ACP for intramolecular loading of the KS  
15 of extension module 1 (Schwecke, T. et al. Proc. Natl.  
Acad. Sci. USA (1995) 92:7839-7843).

The DNA sequences have been disclosed for several Type  
I PKS gene clusters that govern the production of 16-  
membered macrolide polyketides, including the tylosin PKS  
20 from *Streptomyces fradiae* (EP 0 791 655 A2), the  
niddamycin PKS from *Streptomyces caelestis* (Kavakas, S. J.  
et al. J. Bacteriol. (1998) 179:7515-7522) and the  
spiramycin PKS from *Streptomyces ambofaciens* (EP 0791 655  
A2). All of these gene sequences have in common that they  
25 show the loading module of the PKS to differ from the  
loading module of DEBS and of the avermectin PKS in that  
they consist of a domain resembling the KS domains of the  
extension modules, an AT domain and an ACP (Figure 3). The  
additional N-terminal KS-like domain has been named KSq

because it differs in each case from an extension KS by the specific replacement of the active site cysteine residue essential for  $\beta$ -ketoacyl-ACP synthase activity by a glutamine (Q in single letter notation) residue. The function of the KSq domain is unknown (Kavakas, S. J. et al. J. Bacteriol. (1998) 179:7515-7522), but its presence in these PKSs for 16-membered macrolides is surprising because the starter units of tylosin, niddamycin and spiramycin appear to be propionate, acetate and acetate respectively, that is, the same type of starter unit as in DEBS. The AT adjacent to the KSq domain is named here the ATq domain.

When the entire loading module of the tylosin PKS was used to replace the analogous loading module in the spiramycin PKS in *S. ambofaciens* (Kuhstoss et al. Gene (1996) 183:231-236), the nature of the starting unit was stated to be altered from acetate to propionate. Since the role of the KSq domain was not understood, no specific disclosure was made that revealed either the importance of the KSq domain, or the possible utility of these KSq-containing loading modules in ensuring the purity of the polyketide product in respect of the starter unit, even at high levels of macrolide production. The interpretation for their results was stated as: "Therefore we believe that the experiments described here provide strong experimental support for the hypothesis that the AT domains in Type I PKS systems select the appropriate substrate at each step in synthesis" (Kuhstoss et al. Gene (1996) 183:231-236, at p. 235). These authors noted the analogy with the CLF

protein in Type II PKS systems and that the latter protein is thought to be involved in determining the chain length. They state: "KSq may serve a similar function, although it is unclear why such a function would be necessary in the synthesis of these 16-membered polyketides when it is not needed for the synthesis of other complex polyketides such as 6-DEB or rapamycin. In any case the KSq is unlikely to be involved in substrate choice at each step of synthesis." (Kuhstoss et al. Gene (1996) 183:231-236).

It has been shown that when genetic engineering is used to remove the loading module of DEBS, the resulting truncated DEBS in *Sacch. erythraea* continues to produce low levels of erythromycins containing a propionate starter unit (Pereda, A. et al. Microbiology (1995) 144:543-553).

The same publication shows that when in this truncated DEBS the methylmalonyl-CoA -specific AT of extension module 1 was replaced by a malonyl-CoA-specific AT from an extension module of the rapamycin PKS, the products were also low levels of erythromycins containing a propionate starter unit, demonstrating that the origin of the starter units is not decarboxylation of the (methyl)malonyl groups loaded onto the enzyme by the AT of module 1, but from direct acylation of the KS of extension module 1 by propionyl-CoA. This is in contrast to a previous report, using partially purified DEBS1+TE, a truncated bimodular PKS derived from DEBS (Kao, C. M. et al. J. Am. Chem. Soc. (1995) 117:9105-9106) and functionally equivalent to DEBS1-TE (Brown, M. J. B. et al., J. Chem. Soc. Chem. Commun. (1995) 1517-1518;

Cortés, J. et al. Science (1991) 2523:675-679), which stated that the origins of the starter units for DEBS can include methylmalonate units which are loaded onto module 1 and are decarboxylated by the KS of module 1 (Pieper, R. et al. Biochemistry (1997) 36:1846-1851). It has now been found that when the DEBS1-TE protein is fully purified from extracts of recombinant *Sacch. erythraea* it contains no such specific decarboxylase activity (Weissmann, K. et al. (1998) Biochemistry, 37, 11012-11017), further confirming that starter units do not in fact arise from decarboxylation of extension units mediated by the KS of extension module 1 .

It is known that the DEBS loading module has a slightly broader specificity than propionate only, and in particular acetate starter units are used both in vitro and in vivo, when the PKS containing this loading module is part of a PKS that is expressed either in *Sacch. erythraea* the natural host for erythromycin production (see for example Cortés, J. et al. Science (1995) 268:1487-1489), or in an heterologous host such as *S. coelicolor* (Kao, C. M. et al. J. Am. Chem. Soc. (1994) 116:11612-11613; Brown, M. J. B. et al. J. Chem. Soc. Chem. Commun. (1995) 1517-1519). In vitro experiments using purified DEBS1-TE have demonstrated that propionyl-CoA and acetyl-CoA are alternative substrates that efficiently supply propionate and acetate units respectively to the loading module (Wiessmann, K. E. H. et al. Chemistry and Biology (1995)

2:583-589; Pieper, R. et al. J. Am. Chem. Soc. (1995)  
117:11373-11374). The outcome of the competition between  
acetate and propionate starter units is influenced by the  
respective intracellular concentrations of propionyl-CoA  
5 and acetyl-CoA prevailing in the host cell used (see for  
example Kao, C. M. et al. Science (1994) 265:509-512;  
Pereda, A. et al. Microbiology (1995) 144:543-553). It is  
also determined by the level of expression of the host PKS,  
so that as disclosed for example in Pending International  
10 Patent Application number PCT/GB97/01819, when recombinant  
DEBS or another hybrid PKS containing the DEBS loading  
module is over-expressed in *Sacch. erythraea*, the products  
are generally mixtures whose components differ only in the  
presence of either an acetate or a propionate starter unit.

15  
There is a need to develop reliable methods for  
avoiding the formation of mixtures of polyketides with both  
acetate and propionate starter units, and to allow the  
specific incorporation of unusual starter units. It has now  
20 been found, surprisingly, that the role of the loading  
domains in the PKSs for the 16-membered macrolides tylosin,  
niddamycin and spiramycin is different from that of the  
loading domains of the avermectin PKS and of DEBS. It has  
been realised that the KSq domain of the tylosin PKS and  
25 the associated AT domain, which is named here ATq, together  
are responsible for the highly specific production of  
propionate starter units because the ATq is specific for  
the loading of methylmalonyl-CoA and not propionyl-CoA as  
previously thought; and the KSq is responsible for the

highly specific decarboxylation of the enzyme-bound methylmalonate unit to form propionate unit attached to the ACP domain of the loading module and appropriately placed to be transferred to the KS of extension module 1 for the initiation of chain extension. In a like manner the ATq of the spiramycin and niddamycin PKSs, and the adjacent KSq, are responsible for the specific loading of malonate units rather than acetate units as previously believed, and for their subsequent specific decarboxylation to provide acetate starter units for polyketide chain extension.

It has also now been found here that not only the PKSs for the above-mentioned 16-membered macrolides, but also the PKSs for certain 14-membered macrolides particularly the oleandomycin PKS from *Streptomyces antibioticus* (Figure 4) and also the PKSs for certain polyether ionophore polyketides particularly the putative monensin PKS from *Streptomyces cinnamonensis* (Figure 4), possess a loading domain comprising a KSq domain, an ATq domain, and an ACP. In Figure 4 is shown a sequence alignment of the KSq domains and of the adjacent linked ATq domains that have been identified, showing the conserved active site glutamine (Q) residue in the KSq domains, and an arginine residue which is conserved in all extension AT domains and is also completely conserved in ATq domains. This residue is characteristically not arginine in the AT domains of either DEBS or of the avermectin PKS loading modules, where the substrate for the AT is a non-carboxylated acyl-CoA ester (Haydock, S. F. et al. FEBS Letters (1995) 374:246-248) . The abbreviation ATq is used here to simply to



distinguish the AT domains found immediately C-terminal of Ksq from extension ATs, and the label has no other significance.

In one aspect this invention provides a PKS  
5 multienzyme or part thereof, or nucleic acid (generally DNA) encoding it, said multienzyme or part comprising a loading module and a plurality of extension modules for the generation of novel, 14-membered macrolides wherein

(a) the loading module is adapted to load a malonyl  
10 residue and then to effect decarboxylation of the loaded residue to provide an acetyl residue for transfer to an extension module; and

(b) the extension modules, or at least one thereof (preferably at least the one adjacent the loading module),  
15 are not naturally associated with a loading module that effects decarboxylation of an optionally substituted malonyl residue.

Generally the loading module will also include an ACP (acyl carrier protein) domain.

20 Preferably the decarboxylating functionality of the loading module is provided by a KS (ketosynthase)-type domain. Suitably this differs from a KS of a conventional extension module by possessing a glutamine residue in place of the essential cysteine residue in the active site. It  
25 is termed Ksq. It may be "natural" or genetically engineered, e.g. resulting from site-directed mutagenesis of nucleic acid encoding a different KS such as a KS of an extension module.

Alternatively the decarboxylating functionality can be provided by a CLF-type domain of the general type occurring in Type II PKS systems.

Preferably the loading functionality is provided by an  
5 AT (acyltransferase)-type domain which resembles an AT domain of a conventional extension module in having an arginine residue in the active site, which is not the case with the AT domains of loader modules which load acetate or propionate, e.g. in DEBS or avermectin PKS systems. It may  
10 be termed Atq. Once again, it may be "natural" or genetically engineered, e.g. by mutagenesis of an AT of an extension module.

Usually the loading module will be of the form:

Ksq-ATq-ACP

15 where ACP is acyl carrier protein.

In another aspect the invention provides a method of synthesising novel, 14-membered polyketides having substantially exclusively a desired acetate starter unit by providing a PKS multienzyme incorporating a loading module  
20 as defined above which specifically provides the desired acetate starter unit. This may comprise providing nucleic acid encoding the multienzyme and introducing it into an organism where it can be expressed.

In further aspects the invention provides vectors and  
25 transformant organisms and cultures containing nucleic acid encoding the multienzyme. A preferred embodiment is a culture which produces a 14-membered polyketide having a desired acetate starter unit characterised by the substantial absence of polyketides with different starter  
30 units. Thus, for example, C13-methyl-erythromycin can be

produced substantially free from natural analogues resulting from the incorporation of propionate starter units.

It is particularly useful to provide a loading module of the type KSq - ATq-ACP for a PKS gene assembly which produces a 14-membered macrolide in order to prepare a 14-membered macrolide which contains exclusively or almost exclusively an acetate starter unit, even when such PKS gene assembly is expressed at high levels in an actinomycete host cell. Particularly suitable PKSs for this purpose are the components of PKSs for the biosynthesis of erythromycin, methymycin, oleandomycin, tylosin, spiramycin, midecamycin, and niddamycin for all of which the gene and modular organisation is known at least in part. Particularly suitable sources of the genes encoding a loading module of the type KSq - ATq-ACP are the loading modules of oleandomycin, spiramycin, niddamycin, methymycin and monensin which are specific for the loading of malonate units which are then decarboxylated to acetate starter units.

In the loading module of the type KSq - ATq-ACP the domains or portions of them may be derived from the same or from different sources, and comprise either natural or engineered domains. For example the ATq domain can be replaced by an AT domain derived from any extension module of a Type I PKS, having specificity for loading of malonate units, so long as the KSq domain is chosen to have a matching specificity towards malonate units.

Alternatively, the KSq domain in the loading module provided of the type KSq - ATq-ACP may be substituted by

the CLF polypeptide of a Type II PKS. It is now apparent that in contrast to its previous identification as a factor uniquely determining chain length, the CLF, in addition to any other activities that it may possess, is the analogue  
5 of the KSq domain and can act as a decarboxylase towards bound malonate units.

The appreciation that the CLF domain of Type II PKS's has decarboxylating activity has led us to devise useful interventions in Type II systems, e.g. to enhance the  
10 yields obtainable in some fermentations. Many high-yielding industrial fermentations tend to give mixtures, owing to the incorporation of undesired starters. This is particularly the case in systems which have auxiliary genes for generating unusual starters. CLF genes may act to  
15 produce undesired acyl species, leading to products incorporating the undesired acyl units.

For example the production of oxytetracycline involves an unusual malonamido starter. However the undesired activity of a CLF domain causes some decarboxylation,  
20 leading to the incorporation of acetyl instead. Daunomycin synthesis likewise involves an unusual starter which is liable to the "parasitic" activity of a CLF domain.

The active site (for decarboxylation) of a CLF domain generally includes a glutamine residue. We find that the  
25 decarboxylating activity of the domain can be removed by a mutation by which the Gln residue is converted into (for example) Ala.

Thus in a further aspect the invention provides a

system and process for synthesis of a type II (aromatic) polyketide, in which a gln residue of a CLF domain of the type II PKS is mutated to suppress decarboxylation activity. Techniques of site-specific mutagenesis by which this can be achieved are by now well known to those skilled in the art.

The loading module of the type KSq - ATq-ACP may be linked to a hybrid PKS produced for example as in PCT/GB97/01819 and PCT/GB97/01810. It is particularly useful to link such a loading module to gene assemblies that encode hybrid PKSs that produce novel derivatives of 14-membered macrolides as described for example in PCT/GB97/01819 and PCT/GB97/01810.

The invention further provides such PKS assemblies furnished with a loading module of the type KSq - ATq-ACP, vectors containing such assemblies, and transformant organisms that can express them. Transformant organisms may harbour recombinant plasmids, or the plasmids may integrate. A plasmid with an *int* sequence will integrate into a specific attachment site (*att*) of the host's chromosome. Transformant organisms may be capable of modifying the initial products, eg by carrying out all or some of the biosynthetic modifications normal in the production of erythromycins (as shown in Figure 5) and for other polyketides. Use may be made of mutant organisms such that some of the normal pathways are blocked, e.g. to produce products without one or more "natural" hydroxy-groups or sugar groups. The invention further provides novel polyketides as producible, directly or indirectly, by

transformant organisms. This includes polyketides which have undergone enzymatic modification.

In a further aspect the invention provides both previously-obtained 14-membered ring macrolides and novel  
5 14-membered ring macrolides in a purer form with respect to the nature of the acetate starter unit, than was hitherto possible. These include 14-membered ring macrolides which are either "natural" or may differ from the corresponding "natural" compound:

10

a) in the oxidation state of one or more of the ketide units (ie selection of alternatives from the group: -CO-, -CH(OH)-, alkene -CH-, and -CH<sub>2</sub>- ) where the stereochemistry of any -CH(OH)- is also independently selectable;

15

b) in the absence of a "natural" methyl side-chain; or

c) in the stereochemistry of "natural" methyl; and/or ring substituents other than methyl.

20

It is also possible to prepare derivatives of 14-membered ring macrolides having the differences from the natural product identified in two or more of items a) to c) above.

25

Derivatives of any of the afore-mentioned polyketides which have undergone further processing by non-PKS enzymes, eg one or more of hydroxylation, epoxidation, glycosylation and methylation may also be prepared.

30

The present invention provides a novel method of obtaining both known and novel complex 14-membered macrolides having an acetate starter unit substantively

free of products differing only in having a propionate starter unit.

Suitable plasmid vectors and genetically engineered cells suitable for expression of PKS genes incorporating an altered loading module are those described in PCT/GB97/01819 as being suitable for expression of hybrid PKS genes of Type I. Examples of effective hosts are *Saccharopolyspora erythraea*, *Streptomyces coelicolor*, *Streptomyces avermitilis*, *Streptomyces griseofuscus*, *Streptomyces cinnamonensis*, *Streptomyces fradiae*, *Streptomyces longisporoflavus*, *Streptomyces hygroscopicus*, *Micromonospora griseorubida*, *Streptomyces lasaliensis*, *Streptomyces venezuelae*, *Streptomyces antibioticus*, *Streptomyces lividans*, *Streptomyces rimosus*, *Streptomyces albus*, *Amycolatopsis mediterranei*, and *Streptomyces tsukubaensis*. These include hosts in which SCP2\*-derived plasmids are known to replicate autonomously, such as for example *S. coelicolor*, *S. avermitilis* and *S. griseofuscus*; and other hosts such as *Saccharopolyspora erythraea* in which SCP2\*-derived plasmids become integrated into the chromosome through homologous recombination between sequences on the plasmid insert and on the chromosome; and all such vectors which are integratively transformed by suicide plasmid vectors.

25

Although some 13-methyl erythromycins (also known as 15-norerythromycins) have been reported previously (Kibwage et

al., J. Antibiotics, 40, 1-6, 1987; Weber & McAlpine, U.S. Patent 5,141,926), these have been confined to 15-norerythromycin C, and 6-deoxy-15-norerythromycins B and D. Moreover, not only have these 15-norerythromycins been found as extremely minor components co-expressed with high levels of "natural" erythromycins (13-ethyl erythromycins), but the 13-methyl counterparts (15-norerythromycins A and B) to the most desirable and biologically-active "natural" erythromycins (erythromycin A and B) have never been previously isolated. Chemical modification of "natural" erythromycins has proven to be an extremely effective means for enhancing the bioefficacy of the "natural" molecules. Thus, it would be envisaged that chemical modification of novel erythromycins would similarly produce compounds with desirable and enhanced bioefficacies. PCT/GB97/01819 describes in general terms the production of novel polyketides through recombinant DNA technologies, and the use of these technologies to generate novel erythromycins, many of which have different starter units to the propionate starter unit characteristic of the "natural" erythromycins, are described in pending International Patent Application PCT/GB97/01810, Some chemical modification of these novel erythromycins are also described in co-pending International Patent Applications PCT/IB98/02100 and PCT/IB98/02099. However, it is clear that the ability to produce novel erythromycins at good expression levels and in the substantial absence of novel or natural



erythromycins with different starter units is essential to facilitate the ability to achieve a wide range of chemical modifications to such novel erythromycins. The enhanced ability to produce polyketides at good expression levels and in the substantial absence of polyketides with different starter units has been described in this application are family members, and we now describe the ability to produce 13-methyl erythromycins at good expression levels and in the substantial absence of erythromycins with different starter units. The use of this technology has now permitted the preparation of large amounts of 13-methyl erythromycins which for the first time has permitted us to carry out a wide range of chemical modifications which had only been previously possible starting from the "natural" erythromycins.

Some embodiments of the invention will now be described with reference to the accompanying drawings in which:

Fig 1 is a diagram showing the functioning of 6-deoxyerythronolide B synthase (DEBS), a modular PKS producing 6-deoxyerythronolide B (6-DEB) a precursor of erythromycin A.

Fig 2 gives the amino acid sequence comparison of the KS domains and the CLF domains of representative Type II PKS gene clusters. The active site Cysteine (C) of the KS domains is arrowed in the Figure and aligns with the Glutamine (Q) or glutamic acid (E) of the CLF domains. The

abbreviations used, and the relevant Genbank/EMBL accession numbers are: GRA: granaticin from *Streptomyces violaceoruber* (X63449); HIR: unknown polyketide from *Saccharopolyspora hirsuta* (M98258); ACT, actinorhodin from *Streptomyces coelicolor* (X63449); CIN: unknown polyketide from *Streptomyces cinnamonensis* (Z11511); VNZ: jadomycin from *Streptomyces venezuelae* (L33245); NOG: anthracyclines from *Streptomyces nogalater* (Z48262); TCM: tetracenomycin from *S. glaucescens* (M80674); DAU: daunomycin from *Streptomyces* sp. C5 (L34880); PEU, doxorubicin from *Streptomyces peucetius* (L35560); WHI: WhiE spore pigment from *Streptomyces coelicolor* (X55942).

Fig 3 shows the gene organisation of the PKSs for three 16-membered ring macrolides, tylosin, spiramycin and niddamycin.

Fig 4 shows the amino acid sequence alignment of KSq-ATq loading didomains of the PKSs for niddamycin, platenolide(spiramycin), monensin, oleandomycin and tylosin. The sequences for the monensin and oleandomycin loading didomains have not been previously disclosed.

Fig. 5 The enzymatic steps that convert 6-deoxyerythronolide B into erythromycin A in *Saccharopolyspora erythraea*

Fig. 6 is a diagram showing the construction of plasmid pJLK117.

Fig. 7 shows the structures of two oligonucleotides. The present invention will now be illustrated, but is

not intended to be limited, by means of some examples.

5 All NMR spectra were measured in CDCl<sub>3</sub> using a Bruker 500MHz DMX spectrometer unless otherwise indicated and peak positions are expressed in parts per million (ppm) downfield from tetramethylsilane. The atom number shown in the NMR structure is not representative of standard nomenclature, but correlates NMR data to that particular example.

10 HPLC methods

Method A

	Column	Waters Symmetry 5_ C18 2.1mm X 150 mm
	Flow	0.29 ml/min
15	Mobile phase	Gradient: A:B (22:78) to A:B (38:62) over 12 minutes, then to A:B (80:20) by minute 15. Maintain for 1 minute. Re-equilibrate before next sample. Where A = acetonitrile and B = 0.01M ammonium acetate in 10% acetonitrile and 0.02% TFA
20	Instrument	Acquired with Hewlett-Packard 1050 liquid chromatograph interfaced to a VG Platform II mass spectrometer equipped with an APCI source
25		

Method B

	Column	Waters Symmetry 5_ C18 2.1mm X 150 mm
	Flow	0.29 ml/min
30	Mobile phase	Gradient: 28:72 acetonitrile:10mM NH <sub>4</sub> OAc to 50:50 in 18 minutes. 50:50 until 25 minutes. Back to 28:72, re-equilibrate for 7 minutes
	Instrument	Acquired with Hewlett Packard 1100 LC/MS with APCI source

35

Tap Water medium

	glucose	5g/liter
	tryptone	5g/liter
40	yeast extract	2.5g/liter
	EDTA	36mg/liter
	Tap water to 1L total volume	

ERY - P medium

45	dextrose	50g/liter
	Nutrisoy <sup>TM</sup> flour	30g/liter
	(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	3g/liter
	NaCl	5g/liter
	CaCO <sub>3</sub>	6g/liter
50	Tap water to 1L total volume	
	pH adjusted to 7.0	

Example 1Construction of the Recombinant Vector pPFL43

5           Plasmid pCJR24 was prepared as described in  
PCT/GB97/01819. pPFL43 is a pCJR24-based plasmid  
containing the gene encoding a hybrid polyketide synthase  
that contains the putative monensin PKS loading module  
(isolated from *S. cinnamomensis*) the DEBS extension  
10 modules 1 and 2 and the chain-terminating thioesterase.  
Plasmid pPFL43 was constructed as follows:

The following synthetic oligonucleotides: 5'-  
CCATATGGCCGCATCCGCGTCAGCGT-3' and 5'-  
15 GGCTAGCGGGTCCTCGTCCGTGCCGAGGTCA-3'  
are used to amplify the DNA encoding the putative  
monensin-producing loading module using a cosmid that  
contains the 5' end of the putative monensin-producing  
PKS genes from *S. cinnamomensis* or chromosomal DNA of *S.*  
20 *cinnamomensis* as template. The PCR product of 3.3 kbp is  
purified by gel electrophoresis, treated with T4  
polynucleotide kinase and ligated to plasmid pUC18, which  
has been linearised by digestion with *Sma* I and then  
treated with alkaline phosphatase. The ligation mixture  
25 was used to transform electrocompetent *E.coli* DH10B cells  
and individual clones were checked for the desired  
plasmid pPFL40. Plasmid pPFL40 was identified by  
restriction pattern and sequence analysis.

Plasmid pHD30His is a derivative of pNEWAVETE (PCT/GB97/01810) which contains the avermectin loading module, erythromycin extension modules 1 and 2 and the ery thioesterase domain. Plasmid pNEWAVETE was cut with

5 EcoRI and HindIII and a synthetic oligonucleotide linker was inserted that encodes the addition of a C-terminal polyhistidine tail to the polypeptide. The following oligonucleotides:

5'-AATTCACATCACCATCACCATCACTAGTAGGAGGTCTGGCCATCTAGA-3'

10 and

5'-AGCTTCTAGATGGCCAGACCTCCTACTAGTGATGGTGATGGTGATGTG-3'

were annealed together and the duplex was ligated to EcoRI- and HindIII-cut pNEWAVETE. The resulting plasmid was cut with NdeI and XbaI and ligated into plasmid

15 pCJR24 that had been previously cut with same two enzymes, to produce plasmid pND30His.

Plasmid pPFL40 was digested with Nde I and Nhe I and the 3.3 kbp fragment was purified by gel electrophoresis and ligated to pND30-His previously digested with Nde I

20 and Nhe I and treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent E.coli DH10B cells and individual clones were checked for the desired plasmid pPFL43. Plasmid pPFL43 was identified by restriction analysis.

25

## Example 2

### Construction of *S. erythraea* NRRL2338/pPFL43

Plasmid pPFL43 was used to transform *S. erythraea* NRRL2338 protoplasts. Thiostrepton resistant colonies were selected in R2T20 medium containing 10 µg/ml of thiostrepton. Several clones were tested for the presence of pPFL43 integrated into the chromosome by Southern blot hybridisation of their genomic DNA with DIG-labelled DNA containing the *mon* PKS fragment encoding for the loading module. A clone with an integrated copy of pPFL43 was selected in this way.

10 Example 3

Production of 13-methyl-erythromycin A and B using *Sacch. erythraea* NRRL 2338/pPFL43

The culture *Saccharopolyspora erythraea* NRRL2338(pPFL43), constructed with the wild-type loading domain displaced by a monensin loader-D1TE DNA insert, produced as described in Example 2, was inoculated into 30ml Tap Water medium with 50 µg/ml thiostrepton in a 300ml Erlenmeyer flask. After three days incubation at 29°C, this flask was used to inoculate 300 ml of ERY-P medium in a 300 ml flask. The broth was incubated at 29°C at 200 rpm for 6 days. After this time, the whole broth was adjusted to pH 8.5 with NaOH, then extracted with equal volume of ethyl acetate. The ethyl acetate extract was evaporated to dryness at 45°C under a nitrogen stream using a Zymark TurboVap LV Evaporator, then reconstituted in 0.0625 volumes methanol to concentrate the extract 16-fold. The structures of the products were confirmed by LC/MS, Method A. A 4.0 min retention time peak was observed as the major component, with *m/z* value of 720 (M+H)<sup>+</sup>, required for 13-methyl-erythromycin A. A second peak was observed with a retention time of 6.4 min and with *m/z* value of 704 (M+H)<sup>+</sup>, required for 13-methyl-erythromycin B.

25 Example 4

35 Production and Recovery of 13-methyl-erythromycin A and B using *Sacch. erythraea* NRRL-2338 (pPFL43) at 8L scale

*Saccharopolyspora erythraea* NRRL2338 (pPFL43) was inoculated into 1000mls Tap Water medium with 50 µg/ml

thiostrepton in a 2.8l Fernbach flask. After three days incubation at 29°C, this flask was used to inoculate 8l of ERY-P medium in a 14l Microferm fermentor jar (New Brunswick Scientific Co., Inc., Edison, NJ). The broth was incubated at 28°C with an aeration rate of 8l/min, stirring at 800 rpm and with pH maintained between 6.9 and 7.3 with NaOH or H<sub>2</sub>SO<sub>4</sub> (15%). Water was added to maintain volume at the 24 hour volume level. The fermentation was continued for 167 hours. After this time, presence of 13-methyl-erythromycin A and B were confirmed by adjusting a broth sample from the fermentor to pH 8.5 with NaOH, then extracting with equal volume of ethyl acetate. The ethyl acetate extract was evaporated to dryness at 45°C under a nitrogen stream using a Zymark TurboVap LV Evaporator, then reconstituted in 0.25 volumes methanol to concentrate the extract 4-fold. The structures of the products were confirmed by LC/MS, Method A. A 4.1 min retention time peak was observed as the major component, with *m/z* value of 720 (M+H)<sup>+</sup>, required for 13-methyl-erythromycin A. A second peak was observed with a retention time of 6.6 min and with *m/z* value of 704 (M+H)<sup>+</sup>, required for 13-methyl-erythromycin B.

About 35 liters of broth containing approximately 2.8 grams of 13-methyl-erythromycin A were processed for recovery of product. Broth was filtered through a pilot sized Ceraflo ceramic unit and loaded onto a 500ml XAD-16 resin column. The product was eluted using 100% methanol. A 175ml CG-161 adsorption column was prepared and equilibrated with 20% methanol/water. A portion of the product solution was adjusted to 20% methanol and loaded onto the column, no breakthrough of product was observed. Washing of the column with up to 40% methanol/water failed at removing any significant level of impurities. Elution with 50% methanol/water achieved chromatographic separation of the product from the two major impurities, 13-methyl-erythromycin B and a degradation product, 13-methyl-dehydroerythromycin A. The purest cuts were combined and reduced in volume by approximately 75% using evaporation to achieve <10% methanol concentration. To enhance 13-methyl-erythromycin A extraction, solid sodium bicarbonate was added until a total concentration of 250mM was obtained. The aqueous product layer was extracted 2x with methylene chloride, using one-half the total volume each time. The volume was reduced to light yellow solids by evaporation. The 13-methyl-erythromycin A was purified by dissolving the crude crystals into methylene chloride at ambient temperature and diluting to 15% methylene chloride with hexane. The cloudy solution is placed at -10°C for ~30 minutes when the liquid is decanted to a 2<sup>nd</sup> flask, leaving the majority of impurities behind as an oil. The flask is left overnight at -10°C, followed by filtration of off-white 13-methyl-erythromycin A crystals the next day. Approximately 300 milligrams of 13-methyl-erythromycin A were isolated from the partial work-up of the 35l broth volume.

Approximately 100 grams of evaporated mother liquor were utilized further to isolate 13-methyl-erythromycin B. Residual 13-methyl-erythromycin A was removed with repetitive extraction of the initial sample with aqueous acetic acid (pH 5). The subsequent methylene chloride layer was chromatographed on 700 g of silica gel using 20% methanol in methylene chloride. The 13-methyl-erythromycin B enriched fractions, as determined by LC/MS, were combined and evaporated to yield ~11.0 grams of dark oil. The oil was dissolved in a minimal amount of methanol and loaded onto 500 ml of Amberchrom CG-161 resin. The 13-methyl-erythromycin B was eluted at 2 bed volumes per hour with 40% methanol in deionized water. One bed volume fractions were collected and assayed by LC/MS. Fractions 42 through 62 were combined, diluted to ~20% methanol with deionized water, and neutralized to pH 7.5 with sodium bicarbonate. The resulting solution was extracted once with 4l of methylene chloride, concentrated to ~500 ml, and dried over anhydrous magnesium

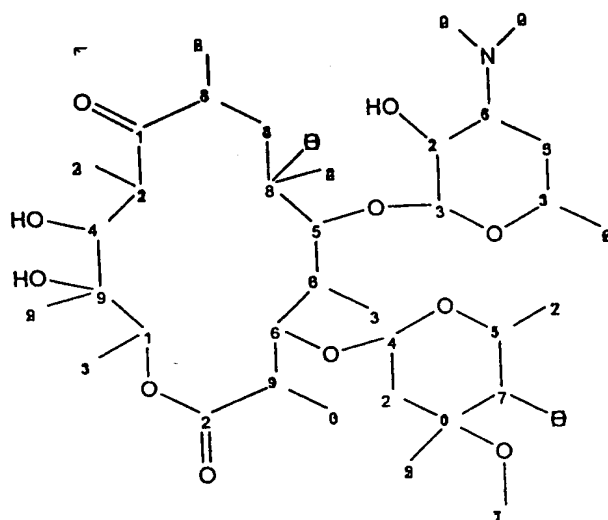
sulfate. After removal of the  $MgSO_4$  by filtration the filtrate was evaporated to give ~110 mg of light brown solids. The 110 mg of crude 13-methyl-erythromycin B was dissolved in ~ 3.0 milliliters of HPLC grade acetonitrile and loaded onto a 20cm x 20cm, 2mm thick, silica gel preparative thin layer chromatography (PTLC) plate. The plate was developed with 60:40 methanol:acetonitrile. The desired portion of silica from the PTLC plate (iodine visualisation) was removed and extracted with HPLC grade acetone. The acetone extract was evaporated to give 12.1 mg of clear solid.

Identification of the 13-methyl-erythromycin A and 13-methyl-erythromycin B samples were confirmed by mass spectroscopy (LC/MS Method B) and NMR spectroscopy. The 13-methyl-erythromycin A sample peak had a 4.7 min retention time, with  $m/z$  value of 720  $(M+H)^+$ , required for 13-methyl-erythromycin A. The 13-methyl-erythromycin B sample peak had a 7.6 min retention time, with  $m/z$  value of 704  $(M+H)^+$ , required for 13-methyl-erythromycin B.

NMR, 13-methyl-erythromycin A:



31

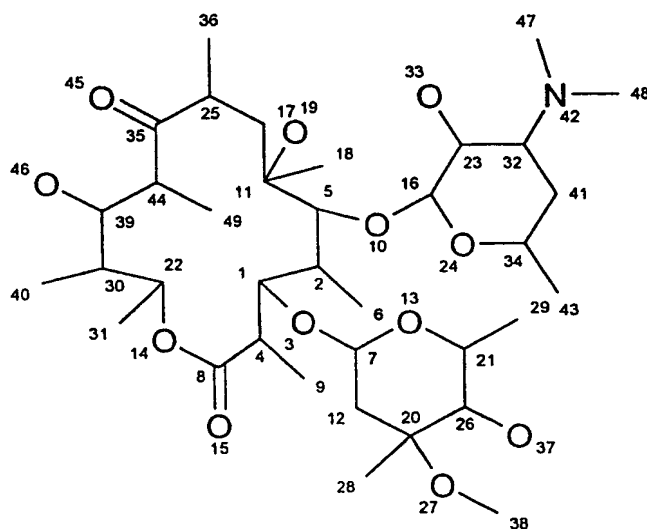


	#	<sup>13</sup> C - ppm	#H	<sup>1</sup> H - ppm
5	1	221.91	0	
	2	175.99	0	
	3	103.63	1	4.45
	4	96.81	1	4.88
10	5	83.76	1	3.60
	6	79.86	1	4.10
	7	78.36	1	3.05
	8	75.50	0	
	9	74.87	0	
15	10	73.07	0	
	11	72.25	1	5.19
	12	71.25	1	3.26
	13	69.53	1	3.53
	14	69.24	1	3.97
20	15	66.16	1	4.06
	16	65.96	1	2.48
	17	49.96	3	3.36
	18	45.36	1	2.79
	19	45.07	1	2.81
25	20	40.73	3	2.32
	21	39.00	1	3.15
	22	35.30	2	2.42/1.61
	24	27.20	3	1.50
	25	21.92	3	1.28
30	26	21.82	3	1.27
	27	18.99	3	1.32
	28	18.60	3	1.22
	29	16.07	3	1.19
	30	15.08	3	1.19
35	31	14.23	3	1.26
	32	12.12	3	1.19

33	9.60	3	1.15
34	39.00	2	1.98/1.75
35	28.90	2	1.72/1.27
36	40.94	1	2.05

5

NMR, 13-methyl-erythromycin B:



10

#	<sup>13</sup> C - PPM	#H attached	<sup>1</sup> H - PPM
1	80.50	1	4.15
2	40.62	1	2.15
4	45.17	1	2.84
5	84.08	1	3.62
6	9.86	3	1.18
7	97.26	1	4.88
8	176.48	0	
9	15.25	3	1.22
11	75.98	0	
12	35.43	2	2.42/1.61
16	103.75	1	4.46
17	38.77	2	2.09/1.72
18	27.67	3	1.51
20	73.09	0	
21	66.20	1	4.06
22	70.27	1	5.58
23	71.24	1	3.28
25	45.49	1	2.81

15

20

25

5	26	78.29	1	3.06
	28	21.91	3	1.28
	29	19.03	3	1.33
	30	41.61	1	1.65
	31	18.73	3	1.29
10	32	65.94	1	2.53
	34	69.52	1	3.55
	35	219.92	0	
	36	19.03	3	1.21
	38	49.97	3	3.36
15	39	70.17	1	3.88
	40	9.27	3	0.95
	41	29.12	2	1.73/1.28
	43	21.80	3	1.27
	44	39.87	1	3.07
20	47	40.74	3	2.35
	48	40.74	3	2.35
	49	9.62	3	1.04

**Example 5****Construction of plasmid pPFL35**

Plasmid pPFL35 is a pCJR24-based plasmid containing a PKS gene comprising a loading module, the first and second extension modules of DEBS and the chain terminating thioesterase. The loading module comprises the KS<sub>q</sub> domain DNA from the loading module of the oleandomycin PKS fused to the malonyl-CoA-specific AT of module 2 of the rapamycin PKS, in turn linked to the DEBS loading domain ACP. Plasmid pPFL35 was constructed via several intermediate plasmids as follows:

A 411 bp DNA segment of the *eryAI* gene from *S. erythraea* extending from nucleotide 1279 to nucleotide 1690 (Donadio, S. et al., Science (1991) 2523:675-679) was amplified by PCR using the following synthetic oligonucleotide primers:-

5'-TGGACCGCCGCAATTGCCTAGGCGGGCCGAACCCGGCT-3' and

5'-CCTGCAGGCCATCGCGACGACCGCGACCGGTTCCGCC-3'

The DNA from a plasmid designated pKSW, derived from pT7-7 and DEBS1-TE in which new *Pst* I and *Hind*III sites had been introduced to flank the KS<sub>1</sub> of the first extension module, was used as a template. The 441 bp PCR product was treated with T4 polynucleotide kinase and ligated to plasmid pUC18, which had been linearised by digestion with *Sma* I and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent *E. coli* DH10B cells and individual clones were checked for the desired plasmid, pPFL26. The new *Mfe* I/*Avr* II sites

bordering the insert are adjacent to the *Eco* RI site in the polylinker of pUC18. Plasmid pPFL26 was identified by restriction pattern and sequence analysis.

An *Mfe* I restriction site is located 112 bp from the  
5 5' end of the DNA encoding the propionyl-CoA:ACP  
transferase of the loading module of DEBS. Plasmid pKSW was  
digested with *Mfe* I and *Pst* I and ligated with the 411 bp  
insert obtained by digesting plasmid pPFL26 with *Mfe* I and  
*Pst* I. The ligation mixture was used to transform  
10 electrocompetent *E.coli* DH10B cells and individual clones  
were checked for the desired plasmid, pPFL27. Plasmid  
pPFL27 contains a PKS gene comprising the DEBS loading  
module, the first and second extension modules of DEBS and  
the DEBS chain terminating thioesterase. Plasmid pPFL27 was  
15 identified by its restriction pattern.

Plasmid pPFL27 was digested with *Nde* I and *Avr* II and  
ligated to a 4.6kbp insert derived from digesting plasmid  
pMO6 (PCT/GB97/01819) with *Nde* I and *Avr* II. Plasmid pMO6  
contains a PKS gene comprising the DEBS loading module, the  
20 first and second extension modules of DEBS and the DEBS  
chain terminating thioesterase, except that the DNA segment  
encoding the methylmalonate-specific AT within the first  
extension module has been specifically substituted by the  
DNA encoding the malonate-specific AT of module 2 of the  
25 *rap* PKS. The ligation mixture was used to transform  
electrocompetent *E. coli* DH10B cells and individual clones  
were checked for the desired plasmid, pPFL28. Plasmid  
pPFL28 contains a hybrid PKS gene comprising the DEBS

loading module, the malonate-specific AT of module 2 of the rap PKS, the ACP of the DEBS loading module, followed by the first and second extension modules of DEBS and the DEBS chain terminating thioesterase. Plasmid pPFL28 was identified by restriction analysis.

A DNA segment encoding the KSq domain from the *oleAI* gene of *S. antibioticus* extending from nucleotide 1671 to nucleotide 3385 was amplified by PCR using the following synthetic oligonucleotide primers:-

10 5'-CCACATATGCATGTCCCCGGCGAGGAA-3' and  
5'-CCCTGTCCGGAGAAGAGGAAGGCGAGGCCG-3'

and chromosomal DNA from *Streptomyces antibioticus* as a template. The PCR product was treated with T4 polynucleotide kinase and ligated to plasmid pUC18, which

15 had been linearised by digestion with *Sma* I and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent *E. coli* DH10B cells and individual clones were checked for the desired plasmid, pPFL31. The new *Nde* I site bordering the insert is adjacent

20 to the *Eco* RI site of the pUC18 polylinker while the new *Bsp* EI site borders the *Hin* dIII site of the linker region. Plasmid pPFL31 was identified by restriction and sequence analysis.

Plasmid pPFL31 was digested with *Nde* I and *Avr* II and

25 the insert was ligated with plasmid pPFL28 that had been digested with *Nde* I and *Avr* II. The ligation mixture was used to transform electrocompetent *E. coli* DH10B cells and

individual clones were checked for the desired plasmid, pPFL32. Plasmid pPFL32 was identified by restriction analysis.

Plasmid pPFL32 was digested with *Nde* I and *Xba* I and the insert was ligated to plasmid pCJR24, which had been digested with *Nde* I and *Xba* I and purified by gel electrophoresis. The ligation mixture was used to transform electrocompetent *E. coli* DH10B cells and individual clones were checked for the desired plasmid, pPFL35. Plasmid pPFL35 was identified by restriction analysis.

#### Example 6

##### Construction of *S. erythraea* NRRL2338/pPFL35

Plasmid pPFL35 was used to transform *S. erythraea* NRRL2338 protoplasts. Thiostrepton resistant colonies were selected in R2T20 medium (Yamamoto et al.) containing 10 µg/ml of thiostrepton. Several clones were tested for the presence of pPFL35 integrated into the chromosome by Southern blot hybridisation of their genomic DNA with DIG-labelled DNA containing the *rap* PKS fragment encoding for module 2 AT. A clone with an integrated copy of pPFL35 was identified in this way.

#### Example 7

##### Production of 13-methyl-erythromycin A and B using *Sacch. erythraea* NRRL-2338 (pPFL35)

The culture *Saccharopolyspora erythraea* NRRL2338(pPFL35), constructed with the wild-type loading

domain displaced by an oleandomycin KSQ-rapamycin AT2-D1TE DNA insert, prepared as described in Example 6, was inoculated into 30ml Tap Water medium with 50 ug/ml thiostrepton in a 300ml Erlenmeyer flask. After two days incubation at 29°C, this flask was used to inoculate 300 ml of ERY-P medium in a 300 ml flask. The broth was incubated at 29°C at 200 rpm for 6 days. After this time, the whole broth was adjusted to pH 8.5 with NaOH, then extracted with an equal volume of ethyl acetate. The ethyl acetate extract was evaporated to dryness at 45°C under a nitrogen stream using a Zymark TurboVap LV Evaporator, then reconstituted in 0.25 volumes methanol to concentrate the extract 4-fold. The structures of the products were confirmed by LC/MS, Method A. A peak was observed with a retention time of 4.0 min and with an  $m/z$  value of 720  $(M+H)^+$ , required for 13-methyl-erythromycin A ( $C_{36}H_{65}NO_{13}$ ). A second peak was observed with a retention time of 6.4 min and with  $m/z$  value of 704  $(M+H)^+$ , required for 13-methyl-erythromycin B ( $C_{36}H_{65}NO_{12}$ ).

#### 20 Example 8

##### Construction of Recombinant Vector pPFL44

Plasmid pPFL44 is a pCJR24- based plasmid containing the gene encoding a hybrid polyketide synthase that contains the spiramycin PKS loading module, the erythromycin extension modules 1 and 2 and the chain-terminating thioesterase. Plasmid pPFL44 was constructed as follows:

The following synthetic oligonucleotides:

30 5'-CCATATGTCTGGAGAACTCGCGATTTCGCCAGT-3' and  
5'-GGCTAGCGGGTCGTCGTCGTCGCCGCTG-3'

were used to amplify the DNA encoding the spiramycin-producing loading module using chromosomal DNA from the spiramycin producer *S. ambofaciens* prepared according to the method described by Hopwood et al. (1985). The PCR product of 3.3 kbp was purified by gel electrophoresis, treated with T4 polynucleotide kinase and ligated to



plasmid pUC18, which had been linearised by digestion with *Sma* I and then treated with alkaline phosphatase.

The ligation mixture was used to transform electrocompetent *E.coli* DH10B cells and individual clones were checked for the desired plasmid pPFL41. Plasmid pPFL41 was identified by restriction pattern and sequence analysis.

Plasmid pPFL41 was digested with *Nde* I and *Nhe* I and the 3.3 kbp fragment was purified by gel electrophoresis and ligated to pND30 ( a plasmid derived from plasmid pCJR24 having as insert the ave PKS loading module and extension modules 1 and 2 or DEBS and the DEBS thioesterase) (PCTGB97/01810) previously digested with *Nde* I and *Nhe* I and treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent *E.coli* DH10B cells and individual clones checked for the desired plasmid pPFL44. Plasmid pPFL44 was identified by restriction analysis.

20

### Example 9

#### Construction of *Sacch. erythraea* NRRL2338/pPFL44

Plasmid pPFL44 was used to transform *S.erythraea* NRRL2338 protoplasts. Thiostrepton resistant colonies were selected in R2T20 medium containing 10 µg/ml of thiostrepton. Several clones were tested for the presence of pPFL44 integrated into the chromosome by Southern blot hybridisation of their genomic DNA with DIG-labelled DNA

containing the spiramycin PKS fragment encoding for the loading module. A clone with an integrated copy of pPFL44 was identified in this way.

5 **Example 10**

**Production of 13-methyl-erythromycin A and B using *Sacch. erythraea* NRRL-2338 (pPFL44)**

10 The culture *Saccharopolyspora erythraea* NRRL2338 (pPFL44), constructed with the wild-type loading domain displaced by spiramycin loader-D1TE DNA insert, was inoculated into 30ml Tap Water medium with 50 ug/ml thiostrepton in a 300ml Erlenmeyer flask. After three days incubation at 29°C, this flask was used to inoculate 300 ml  
15 of ERY-P medium in a 300 ml flask. The broth was incubated at 29°C at 200 rpm for 6 days. After this time, the whole broth was adjusted to pH 8.5 with NaOH, then extracted with equal volume of ethyl acetate. The ethyl acetate extract was evaporated to dryness at 45°C under a nitrogen stream  
20 using a Zymark TurboVap LV Evaporator, then reconstituted in 0.0625 volumes methanol to concentrate the extract 16-fold. The structures of the products were confirmed by LC/MS, Method A. A 4.0 min retention time peak was observed as the major component, with  $m/z$  value of 720  $(M+H)^+$ ,  
25 required for 13-methyl-erythromycin A ( $C_{36}H_{65}NO_{13}$ ). A second peak was observed with a retention time of 6.4 min and with  $m/z$  value of 704  $(M+H)^+$ , required for 13-methyl-erythromycin B ( $C_{36}H_{65}NO_{12}$ ).

30 **Example 21**

Construction of plasmid pJLK114

Plasmid pJLK114 is a pCJR24 based plasmid containing a  
35 PKS gene comprising the ery loading module, the first and the second extension modules of the ery PKS and the ery chain-terminating thioesterase except that the DNA segment between the end of the acyltransferase and the beginning of the ACP of the second ery extension module

has been substituted by a synthetic oligonucleotide linker containing the recognition sites of the following restriction enzymes: AvrII, BglII, SnaBI, PstI, SpeI, NsiI, Bsu36I and HpaI. It was constructed via several  
5 intermediate plasmids as follows (Figure 6).

#### Construction of plasmid pJLK02

The approximately 1.47 kbp DNA fragment of the eryAI gene  
10 of *S. erythraea* was amplified by PCR using as primers the synthetic oligonucleotides:

5'-TACCTAGGCCGGGCCGACTGGTTCGACCTGCCGGGTT-3' and

5'-ATGTTAACCGGTCGCGCAGGCTCTCCGTCT-3' and plasmid pNTEP2  
(Oliynyk, M. et al., Chemistry and Biology (1996) 3:833-  
15 839; WO98/01546) as template. The PCR product was treated with T4 polynucleotide kinase and then ligated with plasmid pUC18, which had been linearised by digestion with SmaI and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent  
20 *E. coli* DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK02 was identified by its restriction pattern and DNA sequencing.

#### Construction of plasmid pJLK03

25

The approximately 1.12 kbp DNA fragment of the eryAI gene of *S. erythraea* was amplified by PCR using as primers the synthetic oligonucleotides:

5'-ATGTTAACGGGTCTGCCGCGTGCCGAGCGGAC-3' and

5'-CTTCTAGACTATGAATTCCCTCCGCCCAGC-3' and plasmid pNTEPH as template. The PCR product was treated with T4 polynucleotide kinase and then ligated with plasmid pUC18, which had been linearised by digestion with SmaI and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK03 was identified by its restriction pattern and DNA sequencing.

10

#### Construction of plasmid pJLK04

Plasmid pJLK02 was digested with PstI and HpaI and the 1.47 kbp insert was ligated with plasmid pJLK03 which had been digested with PstI and HpaI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK04 was identified by its restriction pattern.

20

#### Construction of plasmid pJLK05

Plasmid pJLK01 (PCT/GB97/01819) was digested with PstI and AvrII and the 460 bp insert was ligated with plasmid pJLK04 which had been digested with PstI and AvrII. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK05 was identified by its restriction pattern.

30

### Construction of plasmid pJLK07

Plasmid pJLK05 was digested with ScaI and XbaI and plasmid pNTEPH was digested with NdeI and ScaI and these  
5 two fragments were ligated with plasmid pCJR24 which had been digested with NdeI and XbaI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK07 was  
10 identified by its restriction pattern.

### Construction of plasmid pJLK114

The two synthetic oligonucleotides Plf and Plb (Figure 7)  
15 were each dissolved in TE-buffer. 10  $\mu$ l of each solution (0.5nmol/ $\mu$ l) were mixed and heated for 2 minutes to 65C and then slowly cooled down to room temperature. Plasmid pJLK07 was digested with AvrII and HpaI and ligated with the annealed oligonucleotides. The ligation mixture was  
20 used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK114 was identified by its restriction pattern.

25 Plasmid pJLK117 is a pCJR24 based plasmid containing a PKS gene comprising the ery loading module, the first and the second extension modules of the ery PKS and the ery chain-terminating thioesterase except that the DNA segment between the end of the acyltransferase and the

beginning of the ACP of the second ery extension module has been substituted by a synthetic oligonucleotide linker containing the recognition sites of the following restriction enzymes. AvrII, BglII, SnaBI, PstI, SpeI, NsiI, Bsu36I and NheI.

It was constructed via several intermediate plasmids as follows (Figure 6).

#### 10 Construction of plasmid pJLK115

Plasmid pJLK114 was digested with NdeI and XbaI and the approximately 9.9 kbp insert was ligated with plasmid pUC18 which had been digested with NdeI and XbaI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK115 was identified by its restriction pattern.

#### 20 Construction of plasmid pJLK116

Plasmid pJLK13 (PCT/GB97/01819) was digested with Bsu36I and XbaI and the 1.1 kbp fragment was ligated with plasmid pJLK115 which had been digested with Bsu36I and XbaI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK116 was identified by its restriction pattern.

### Construction of plasmid pJLK117

Plasmid pJLK116 was digested with NdeI and XbaI and the 9.9 kbp fragment was ligated with plasmid pCJR24 which  
5 had been digested with NdeI and XbaI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK117 was identified by its restriction pattern.

10

### Example 11

#### Construction of plasmid pJLK29

15 Plasmid pJLK29 is a pJLK117 based plasmid except that the DNA fragment encoding the reductive loop of module 10 of the rap PKS has been inserted into the mcs. It was constructed via several intermediate plasmids as follows. (Figure 5)

20

#### Construction of plasmid pJLK121.1

The approximately 2.2 kbp DNA segment of the rapB gene of S. hygroscopicus encoding the reductive loop of module 10  
25 was amplified by PCR using as primers the synthetic oligonucleotides:

5'-TAAGATCTTCCGACGTACGCGTTCCAGC-3' and

5'-ATGCTAGCCACTGCGCCGACGAATCACCGGTGG-3' and as template an approximately 7 kbp fragment, which has been obtained

by digestion of cosmid cos 26 (Schwecke, T. et al. (1995) Proc. Natl. Acad. Sci. USA 92:7839-7843) with ScaI and SphI. The PCR product was treated with T4 polynucleotide kinase and then ligated with plasmid pUC18, which had  
5 been linearised by digestion with SmaI and then treated with alkaline phosphatase. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK121.1 was identified by  
10 its restriction pattern and DNA sequencing.

#### Construction of plasmid pJLK29

Plasmid pJLK121.1 was digested with BglII and NheI and  
15 the 2.2 kbp fragment was ligated with plasmid pJLK117 which had been digested with BglII and NheI. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK29 was  
20 identified by its restriction pattern.

#### Example 24

##### Construction of Plasmid pJLK50

25

The approximately 6.1 kbp DNA segment of the erythromycin PKS gene cluster of *S. erythraea* encoding the DNA fragment from the beginning of the ACP of module 2 to the beginning of the ACP of module 3 was amplified by PCR



using as primers the synthetic oligonucleotides:

5'-TACCTGAGGGACCGGCTAGCGGGTCTGCCGCGTG-3' and

5'-ATGCTAGCCGTTGTGCCGGCTCGCCGGTCCGGTCC-3' and plasmid

pBAM25 (published pBK25 by Best, D J et al. Eur J Biochem

5 (1992) 204: 39-49) as template. The PCR product was  
treated with T4 polynucleotide kinase and then ligated  
with plasmid pUC18, which had been linearised by  
digestion with SmaI and then treated with alkaline  
phosphatase. The ligation mixture was used to transform  
10 electrocompetent *E. coli* DH10B cells and individual  
colonies were checked for their plasmid content. The  
desired plasmid pJLK50 was identified by its restriction  
pattern and DNA sequencing.

#### 15 **Example 25**

Construction of *S. erythraea* strain JLK10

Strain JLK10 is a variant of strain NRRL2338 in which the  
reductive loop of ery module 2 (i.e. the KR domain) is  
20 replaced by the reductive loop of the rapamycin module  
10. It was constructed using plasmid pJLK54 which was  
constructed as follows.

Construction of plasmid pJLK54

25

Plasmid pJLK54 is a pJLK29 based plasmid containing a PKS  
gene comprising the ery loading module, the first, the  
second and the third extension modules of the ery cluster  
and the ery chain-terminating thioesterase except that

the DNA segment between the end of the acyltransferase and the beginning of the ACP of the second ery extension module has been substituted by the equivalent segment of module 10 of the rapamycin PKS.

5 It was constructed as follows.

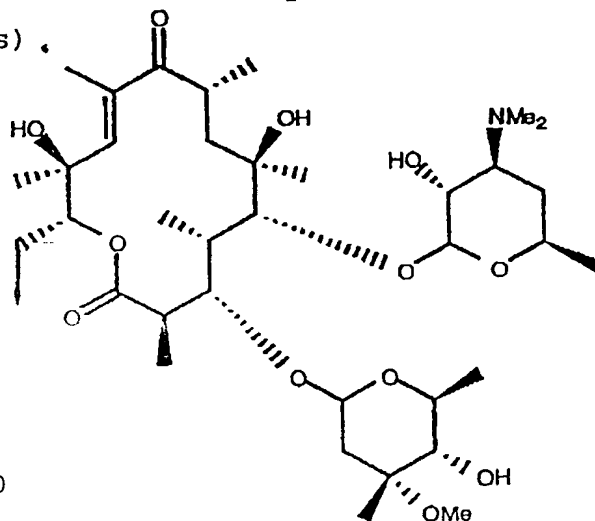
Plasmid pJLK50 was digested with NheI and the 6.1 kbp insert was ligated with plasmid pJLK29 which had been digested with NheI. The ligation mixture was used to  
10 transform electrocompetent *E. coli* DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pJLK54 was identified by its restriction pattern.

15 Use of plasmid pJLK54 for construction of *S. erythraea* NRRL2338/pJLK54 and the production of TKL derivatives

Approximately 5  $\mu$ g plasmid pJLK54 were used to transform protoplasts of *S. erythraea* NRRL2338 and stable  
20 thiostrepton resistant colonies were isolated. From several colonies total DNA is obtained and analysed by Southern blot hybridisation, to confirm that the plasmid has integrated into the TE.

25 Construction of *S. erythraea* strain JLK10 and its use in production of 13-methyl-10,11-dehydro-erythromycin A  
*S. erythraea* strain JLK10 is a mutant of *S. erythraea* NRRL2338 in which the 'reductive loop' of ery module 2 i.e. the ketoreductase domain is substituted by the

'reductive loop' of rapamycin module 10. It was constructed starting from *S. erythraea* NRRL2338 into which plasmid pJLK54 had been integrated. *S. erythraea* NRRL2338/pJLK54 was subjected to several rounds of non-selective growth which resulted in second crossover concomitant with the loss of the integrated plasmid. Clones in which replacement of the erythromycin gene coding for DEBS1 with the mutant version had occurred, were identified by Southern blot hybridisation. One of these was named *S. erythraea* strain JLK10 and was used to inoculate SM3 medium (eryP medium gave similar results), and allowed to grow for seven to ten days at 28-30°C. After this time the broth was centrifuged and the pH of the supernatant adjusted to pH 9. The supernatant was then extracted three times with an equal volume of ethyl acetate and the solvent was removed by evaporation. Products were analysed by HPLC/MS, MS/MS and 1H-NMR. The following macrolide C-13 methyl erythromycin A was identified (accompanied by products of incomplete processing by post-PKS enzymes).



#### Example 26

Construction of plasmid pPFL50

Plasmid pPFL50 is a pPFL43-based plasmid from which a DNA

fragment encoding KR1 (in part), ACP1 and module 2 of the erythromycin PKS and the erythromycin TE, has been removed. It was constructed as follows. Plasmid pPFL43 was digested with SfuI and XbaI to remove a 6.5 kb  
5 fragment. The 5' overhangs were filled in with Klenow fragment DNA Polymerase I and the plasmid was recircularised. The ligation mixture was used to transform electrocompetent *E. coli* DH10B cells and individual colonies were checked for their plasmid  
10 content. The desired plasmid pPFL50 was identified by its restriction pattern.

#### Construction of *S. erythraea* JLK10/pPFL50

Approximately 5 µg plasmid pPFL50 were used to transform  
15 protoplasts of *S. erythraea* strain JLK10 and stable thiostrepton resistant colonies were isolated. From several colonies total DNA was obtained and analysed by Southern blot hybridisation, to confirm that the plasmid had integrated into the homologous chromosomal DNA  
20 region. *S. erythraea* strain JLK10/pPFL50 was used to inoculate SM3 medium containing 5 µg/ml thiostrepton (eryP medium containing 5 µg/ml thiostrepton gave similar results) and allowed to grow for seven to ten days at 28-30°C. After this time the broth was centrifuged and the  
25 pH of the supernatant adjusted to pH 9. The supernatant was then extracted three times with an equal volume of ethyl acetate and the solvent was removed by evaporation. Products were analysed by HPLC/MS, MS/MS and 1H-NMR. The macrolide C-13 methyl 10,11-dehydro-erythromycin A was

identified (accompanied by products of incomplete processing by post-PKS enzymes)

#### Construction of *S. erythraea* NRRL2338/pPFL50

5 Approximately 5 µg plasmid pPFL50 were used to transform protoplasts of *S. erythraea* NRRL2338 and stable thiostrepton resistant colonies were isolated. From several colonies total DNA was obtained and analysed by Southern blot hybridisation, to confirm that the plasmid  
10 had integrated into the homologous region of the chromosomal DNA. *S. erythraea* NRRL2338/pPFL50 was used to inoculate SM3 medium containing 5 µg/ml thiostrepton (eryP medium containing 5 µg/ml thiostrepton gives similar results) and allowed to grow for seven to ten  
15 days at 28-30°C. After this time the broth was centrifuged and the pH of the supernatant adjusted to pH 9.5. The supernatant was then extracted three times with an equal volume of ethyl acetate and the solvent was removed by evaporation. Products were analysed by  
20 HPLC/MS, MS/MS and 1H-NMR. The macrolide C-13 methyl erythromycin A was identified (accompanied by products of incomplete processing by post-PKS enzymes).

#### Construction of plasmid pCB121

25 Plasmid pCB121 is a plasmid containing the monensin loading module and KS of monensin module 1 followed by the erythromycin module 1 AT and part of the erythromycin module 1 KR. It was constructed via several intermediate plasmids as follows.

### Construction of plasmid pPFL45

The approximately 1.8 kbp DNA segment of the monensin PKS gene cluster of *Streptomyces cinnamonensis* encoding part of the ACP of the loading module and KS of module 1 was  
5 amplified by PCR using as primers the synthetic oligonucleotides:

5'-CGTTCCTGAGGTCGCTGGCCCAGGCGTA-3'

5'-CGAAGCTTGACACCGCGGCGGCGCGG-5'

and a cosmid containing the 5' end of the monensin PKS  
10 genes from *S. cinnamonensis* or alternatively chromosomal DNA of *S. cinnamonensis* as template. The PCR product was treated with T4 polynucleotide kinase and then ligated with plasmid pUC18, which had been linearised by digestion with *Sma*I and then treated with alkaline  
15 phosphatase. The ligation mixture was used to transform electrocompetent *E. coli* DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pPFL45 was identified by its restriction pattern.

### Construction of plasmid pPFL47

Plasmid pPFL45 was digested with *Nde*I and *Bsu*36I and the approximately 2.6 kbp fragment was ligated into plasmid pPFL43 which had been digested with *Nde*I and *Bsu*36I. The ligation mixture was used to transform electrocompetent *E. coli* DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pPFL47 was identified by its restriction pattern.

### Construction of plasmid pCB135

Plasmid pCJR24 was digested with *Hind*III, the 5' overhang was filled in with Klenow fragment DNA Polymerase I and  
25 religated. The ligation mixture was used to transform electrocompetent *E. coli* DH10B cells and individual colonies were checked for their plasmid content. The

53

desired plasmid pCB135 was identified by its restriction pattern, lacking the recognition site for HindIII.

#### Construction of plasmid pKSW1

Plasmid pKS1W is a pNTEP2 (GB97/01810)-derived vector containing a DEBS1TE-derived triketide synthase with the unique restriction sites introduced at the limits of KS1.

- 5 Plasmid pKS1W is obtained via several intermediate plasmids as follows.

#### Construction of plasmids pMO09, pMO10 and pMO13

- 10 For the PCR amplification for plasmid pMO09, the following synthetic oligonucleotides were used as mutagenic primers, one containing a MunI site and the other a PstI site:

5' -GCGCGCCAATTGCGTGACATCTCGAT- 3'

and 5' -CCTGCAGGCCATCGCGACGACCGCGACCGGTTGCGCG- 3'

15

For the PCR amplification for plasmid pMO10, the following synthetic oligonucleotides were used as mutagenic primers, one containing a HindIII site and the other an EcoRV site:

- 20 5' -GTCTCAAGCTTCGGCATCAGCGGCACCAA- 3'

and 5' -CGTGCGATATCCCTGCTCGGCGAGCGCA-3'

- 25 For the PCR amplification for plasmid pMO13, the following synthetic oligonucleotides were used as mutagenic primers, one containing a PstI site and the other a HindIII site:

5' -GATGGCCTGCAGGCTGCCCGGCGGTGTGAGCA- 3'

and 5' -GCCGAAGCTTGAGACCCCCGCGCGGTTCGC- 3'

PCR was carried out on pNTEP2 (GB97/01810) as template using Pwo DNA polymerase and one cycle of: 96°C (1min); annealing at 50°C (3min); and extension at 72°C (1min), and 25 cycles of: 96°C (1min); annealing at 50°C (1min);  
5 and extension at 72°C (1min) in the presence of 10% (vol/vol) dimethylsulphoxide. The products were end-repaired and cloned into pUC18 digested with SmaI and the ligation mixture was transformed into E. coli DH 10B. Plasmid DNA was prepared from individual colonies. The  
10 desired plasmids for pMO09 (3.8kbp), pMO10 (3.9 kbp) and pMO13 (4.3 kbp) were identified by their restriction pattern and DNA sequencing.

#### Construction of plasmid pMO11

15 Plasmid pMO13 was digested with HindIII, and the 1.2 kbp insert was cloned into pMO10 which had been digested with HindIII. The ligation mixture was transformed into E. coli DH 10B. The desired plasmid (5.0 kbp) was identified by its restriction pattern and designated  
20 pMO11.

#### Construction of plasmid pMO12

Plasmid pMO09 was digested with PstI, and the 1.6 kbp insert was cloned into pMO11 which had been digested with  
25 PstI. The ligation mixture was transformed into E. coli DH 10B. The desired plasmid (6.6 kbp) was identified by its restriction pattern and designated pMO12.

#### Construction of pKS1W



Plasmid pM012 was digested with MunI and EcoRV, and the 3.9 kbp fragment was cloned into pNTEPH (see below) which had been digested with MunI and EcoRV. The ligation mixture was transformed into E. coli DH 10B. The desired  
5 plasmid (13. kbp) was identified by its restriction pattern and designated pKS1W.

#### Construction of pNTEPH

Plasmid pNTEPH was obtained from pNTEP2 by removing the  
10 HindIII site. pNTEP2 was digested with HindIII, the 5' overhang was filled in with Klenow Fragment DNA Polymerase I and religated. The desired plasmid (13.6 kbp) was identified by its restriction pattern.

#### 15 Construction of plasmid pCB136

Plasmid pKSW1 was digested with NdeI and XbaI and the approximately 11.2 kbp fragment was ligated with plasmid pCB135 which had been digested with NdeI and XbaI. The ligation mixture was used to transform electrocompetent  
20 E. coli DH10B cells and individual colonies were checked for their plasmid content. The desired plasmid pCB136 was identified by its restriction pattern.

#### Construction of plasmid pCB137

25 Plasmid pCB136 was digested with SfuI and XbaI to remove a 6.5 kb fragment, the 5' overhangs were filled in with Klenow Fragment DNA Polymerase I and religated. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked

for their plasmid content. The desired plasmid pCB137 was identified by its restriction pattern.

#### **Construction of plasmid pCB121**

- 5 Plasmid pPFL47 was digested with NdeI and HindIII and the approximately 4.4 kbp insert was ligated with plasmid pCB137 which had been digested with NdeI and HindIII. The ligation mixture was used to transform electrocompetent E. coli DH10B cells and individual colonies were checked  
10 for their plasmid content. The desired plasmid pCB121 was identified by its restriction pattern.

#### **Example**

##### **Construction of *S.erythraea* JLK10/pCB121**

- 15 Approximately 5 µg plasmid pCB121 were used to transform protoplasts of *S. erythraea* JLK10 and stable thiostrepton resistant colonies were isolated. From several colonies total DNA was obtained and analysed by Southern blot hybridisation, to confirm that the plasmid had integrated  
20 into the homologous chromosomal DNA region. *S. erythraea* strain JLK10/pCB121 was used to inoculate SM3 medium containing 5 µg/ml thiostrepton (eryP medium containing 5 µg/ml thiostrepton gave similar results) and allowed to grow for seven to ten days at 28-30°C. After this time  
25 the broth was centrifuged and the pH of the supernatant adjusted to pH 9. The supernatant was then extracted three times with an equal volume of ethyl acetate and the solvent was removed by evaporation. Products were

analysed by HPLC/MS, MS/MS and <sup>1</sup>H-NMR. The macrolide C13-methyl-10,11-dehydro-erythromycin A was identified (accompanied by products of incomplete processing by post-PKS enzymes):

5

### Example

#### Construction of *S. erythraea* NRRL2338/pCB121

Approximately 5 µg plasmid pCB121 were used to transform protoplasts of *S. erythraea* NRRL2338 and stable

10 thiostrepton resistant colonies were isolated. From several colonies total DNA was obtained and analysed by Southern blot hybridisation, to confirm that the plasmid had integrated into the homologous chromosomal DNA region. *S. erythraea* NRRL2338/pPFL50 was used to

15 inoculate SM3 medium containing 5 µg/ml thiostrepton (eryP medium containing 5 µg/ml thiostrepton gave similar results) and allowed to grow for seven to ten days at 28-30°C. After this time the broth was centrifuged and the pH of the supernatant adjusted to pH=9.. The supernatant

20 was then extracted three times with an equal volume of ethyl acetate and the solvent was removed by evaporation. Products were analysed by HPLC/MS, MS/MS and <sup>1</sup>H-NMR. The macrolide C13-erythromycin A was identified (accompanied by products of incomplete processing by post-PKS

25 enzymes):

Although the present invention is illustrated by the examples listed above, they should not be regarded as limiting the scope of the invention. The above

descriptions illustrate for the first time the construction of a Type I PKS gene assembly containing a wholly or partly heterologous KSq-containing loading module and its use to obtain polyketide products of utility as synthetic intermediates or as bioactive materials such as antibiotics. It will readily occur to the person skilled in the art that a wholly or partly heterologous KSq-containing loading module from other PKS gene sets could be used to replace the loading module of DEBS, or indeed into a quite different PKS gene assembly. It will also readily occur to the person skilled in the art that that the additional specificity provided by the more efficient discrimination made between methylmalonyl-CoA and malonyl-CoA by an ATq, followed by specific decarboxylation by a KSq, is preferable to the imperfect discrimination between propionyl-CoA and acetyl-CoA that is a feature of the DEBS loading module and of many other PKS loading modules, in that it maximises the production of a single product rather than a mixture differing from each other in the nature of the starter unit. The avoidance of such mixtures increases yields and avoids the need for tedious and difficult separation procedures.

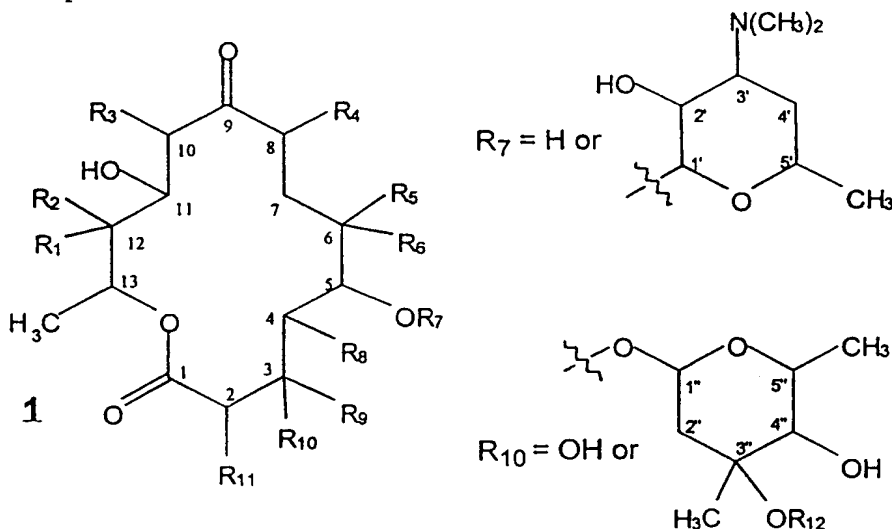
Claims

1. A 14-member macrolide which incorporates an acetate starter unit so that it has a 13-methyl substituent, with the proviso that it is not  
 5 norerythromycin C, 6-deoxy-15-norerythromycin B or 6-deoxy-15-norerythromycin D.

2. 15-norerythromycin A.

10 3. 15-norerythromycin B.

4. A compound of the formula 1:



15

or a pharmaceutically acceptable salt thereof, wherein:

R<sub>1</sub> is H or OH; R<sub>2</sub>-R<sub>4</sub> are each independently H, CH<sub>3</sub>, or CH<sub>2</sub>CH<sub>3</sub>; R<sub>5</sub> is H or OH; and R<sub>6</sub> is H, CH<sub>3</sub>, or CH<sub>2</sub>CH<sub>3</sub>; R<sub>7</sub> is H or desosamine; R<sub>8</sub> is H, CH<sub>3</sub>, or CH<sub>2</sub>CH<sub>3</sub>; R<sub>9</sub> is OH, mycarose  
 20 (R<sub>12</sub> is H), or cladinose (R<sub>12</sub> is CH<sub>3</sub>), R<sub>10</sub> is H; or R<sub>9</sub> = R<sub>10</sub>

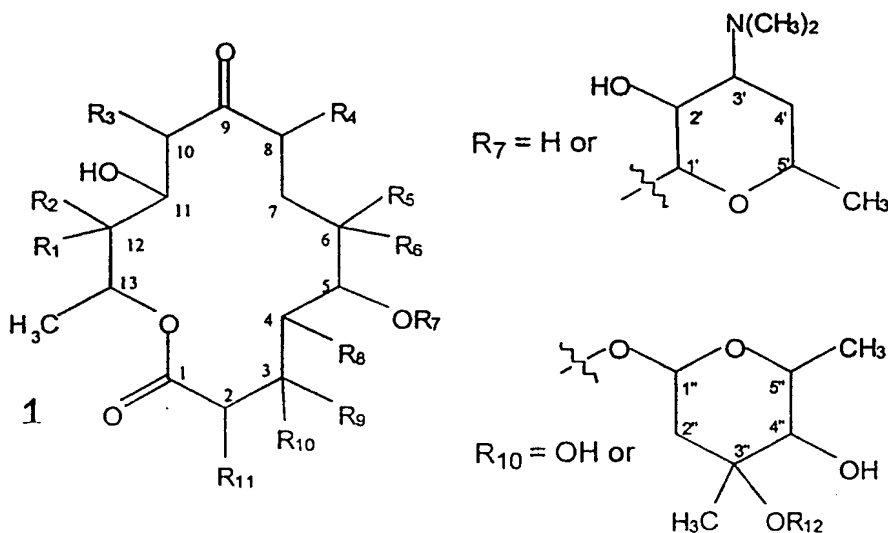
60

= O; and  $R_{11}$  is H,  $\text{CH}_3$ , or  $\text{CH}_2\text{CH}_3$ , with the proviso that when  $R_2$ - $R_4$  are  $\text{CH}_3$ ,  $R_6$  is  $\text{CH}_3$ ,  $R_8$  is  $\text{CH}_3$ , and  $R_{11}$  is  $\text{CH}_3$ , then  $R_1$  and  $R_5$  are not H and  $R_{12}$  is not H; or also when  $R_2$ - $R_4$  are  $\text{CH}_3$ ,  $R_6$  is  $\text{CH}_3$ ,  $R_8$  is  $\text{CH}_3$ , and  $R_{11}$  is  $\text{CH}_3$ , then  $R_1$  and  $R_5$  are not OH and  $R_{12}$  is not H.

5. A compound according to claim 4 wherein  $R_1$  is OH;  $R_2$ - $R_4$  are  $\text{CH}_3$ ;  $R_5$  is OH;  $R_6$  is  $\text{CH}_3$ ,  $R_7$  is desosamine;  $R_8$  is  $\text{CH}_3$ ;  $R_9$  is cladinose ( $R_{12}$  is  $\text{CH}_3$ ); and  $R_{11}$  is  $\text{CH}_3$

10 6. A compound according to claim 4 wherein  $R_1$  is H;  $R_2$ - $R_4$  are  $\text{CH}_3$ ;  $R_5$  is OH;  $R_6$  is  $\text{CH}_3$ ,  $R_7$  is desosamine;  $R_8$  is  $\text{CH}_3$ ;  $R_9$  is cladinose ( $R_{12}$  is  $\text{CH}_3$ ); and  $R_{11}$  is  $\text{CH}_3$ .

7. A process for making compounds of the formula 1:



15

wherein:

$R_1$  is H or OH;  $R_2$ - $R_4$  are each independently H,  $\text{CH}_3$ , or  $\text{CH}_2\text{CH}_3$ ;  $R_5$  is H or OH; and  $R_6$  is H,  $\text{CH}_3$ , or  $\text{CH}_2\text{CH}_3$ ;  $R_7$  is H or desosamine;  $R_8$  is H,  $\text{CH}_3$ , or  $\text{CH}_2\text{CH}_3$ ;  $R_9$  is OH, mycarose

(R<sub>12</sub> is H), or cladinose (R<sub>12</sub> is CH<sub>3</sub>), R<sub>10</sub> is H; or R<sub>9</sub> = R<sub>10</sub> = O; and R<sub>11</sub> is H, CH<sub>3</sub>, or CH<sub>2</sub>CH<sub>3</sub>

8. A process for making compound of the formula 1 as set out in claim 7 wherein R<sub>1</sub> is OH; R<sub>2</sub>-R<sub>4</sub> are CH<sub>3</sub>; R<sub>5</sub> is OH; R<sub>6</sub> is CH<sub>3</sub>, R<sub>7</sub> is desosamine; R<sub>8</sub> is CH<sub>3</sub>; R<sub>9</sub> is cladinose (R<sub>12</sub> is CH<sub>3</sub>); and R<sub>11</sub> is CH<sub>3</sub>

9. A process for making compound of the formula 1 as set out in claim 7 wherein R<sub>1</sub> is H; R<sub>2</sub>-R<sub>4</sub> are CH<sub>3</sub>; R<sub>5</sub> is OH; R<sub>6</sub> is CH<sub>3</sub>, R<sub>7</sub> is desosamine; R<sub>8</sub> is CH<sub>3</sub>; R<sub>9</sub> is cladinose (R<sub>12</sub> is CH<sub>3</sub>); and R<sub>11</sub> is CH<sub>3</sub>

10. A system for producing a 14-membered macrolide incorporating an acetate starter unit, said system comprising DNA encoding and arranged to express a PKS multienzyme which comprises a loading module and a plurality of extension modules; wherein in the expressed multienzyme, said loading module is adapted to load a malonyl residue and then to effect a decarboxylation of the loaded residue to provide an acetate starter unit which is transferred to an adjacent one of said extension modules; and wherein the extension modules, or at least one thereof, are not naturally associated with a loading module that effects decarboxylation.

11. A system according to claim 10 wherein the macrolide is a compound of formula 1 as defined in any of claims 4-9.

12. A system according to claim 10 or 11 wherein said adjacent extension module to which the acetate starter is transferred is not naturally associated with a loading  
5 module that effects decarboxylation.

13. A system according to claim 10, 11 or 12 wherein the decarboxylating functionality of the loading module is provided by a ketosynthase-type domain having a glutamine  
10 residue in the active site.

14. A system according to claim 10, 11 or 12 wherein the decarboxylating functionality of the loading module is provided by a CLF-type domain.  
15

15. A system according to claim 14 wherein the CLF-type domain is substantially as any shown in Fig 2.

16. A system according to any of claims 10-15 wherein  
20 the loading module's loading functionality is provided by an acyltransferase-type domain having an arginine residue in the active site.

17. A system according to any of claims 10-16 wherein  
25 the loading module includes an acyl carrier protein.



18. A system according to any of claims 10-13, 16 or 17 wherein at least the KS<sub>0</sub> domain of said loading module corresponds to the loading module of the PKS multienzyme of oleandomycin, spiramycin, niddamycin, methymycin, or monensin.

19. A PKS multienzyme as expressible by the DNA of the system of any of claims 10-18 or a variant having the ability to synthesise a compound of formula 1.

20. Nucleic acid encoding the PKS multienzyme of claim 19.

21. A vector containing nucleic acid as defined in claim 20.

22. A transformant organism comprising a system according to any of claims 10-18.

23. A process according to claim 7, 8, or 9 which comprises culturing an organism according to claim 22 and recovering a compound of formula 1.

24. A process according to claim 23 wherein said macrolide is a compound of formula 1 as defined in any of claims 4-9.

25. A system, organism or process according to any of  
claims 10-24 wherein the plurality of extension modules  
corresponds to the extension modules of a PKS selected  
5 from erythromycin, narbomycin, pikromycin, lankamycin,  
kujimycin or megalomycin or a mutant or variant thereof  
able to direct synthesis of a macrolide.

# The erythromycin PKS

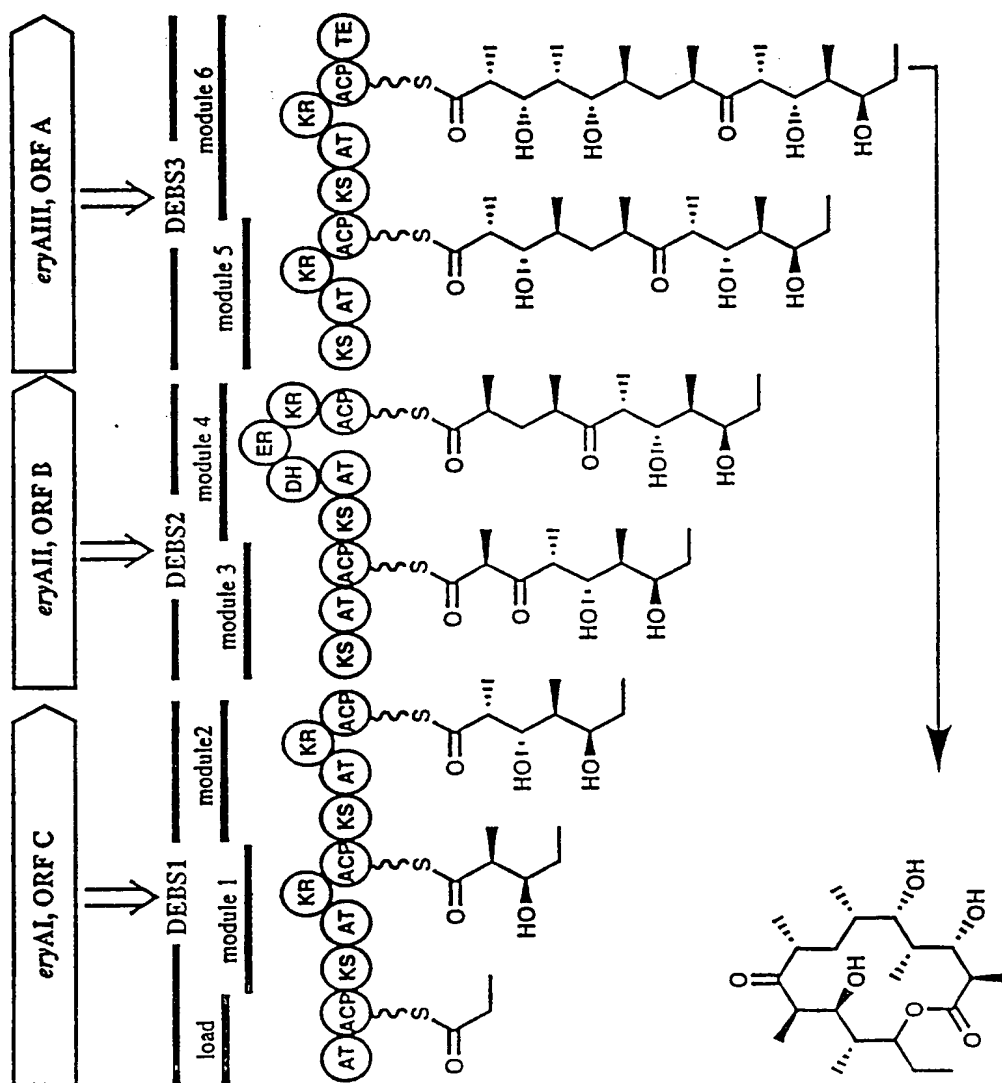


Fig. 1

```

KCLFDAU -----MVTGLGIVAPNGLGVGAIWDVNLNGRNGIGPLR
KCLFPEU MTGTAARTASSQLHASPAGRRGLRGRAVVTGLGIVAPNGLGVGAYWDVNLNGRNGIGPLR
KCLFACT -----MSVLTITGVGVVAPNGLGLAPYWSAVLDGRHGLGPVT
KCLFHIR -----MSTWVTGMGVVAPNGLGADHWATLKGRHGISRLS
KCLFGRA -----MSTPDRRRRAVVTGLSVAAPGGLGTERYWKSLTGENGLAELS
KCLFNOG -----MTAAVVVTGLGVVAPTGLGVREHWSSTVRGASAIGPVT
KCLFTCM -----MSAPAPVVVTGLGIVAPNGTGTEEYWAATLAGKSGIDVIQ
KCLFCIN -----MTP-VAVTGMGLAAPNGLGRPTTGRPPWAPRAASAAT
KCLFVNZ -----MSASVVVTGLGVAAPNGLGREDFWASTLGGKSGIGPLT
KCLFWHIE -----MSGPQRTGTGGSSRRRAVVTGLGVLSPHGTGVEAHWKAVADGTSSSLGPVT
KSGRA -----MTRRVVITGVGVVAPGGSGTKEFWDLLTAGRTATRPIS
KSHIR -----MTRRVVITGVGVVAPGGGLAKNFWELLTSGRTATRRIS
KSACT -----MKRRVVTGVGVVAPGGNGTRQFWELLTSGRTATRRIS
KSCIN -----MTQRRVAITGIEVLAPGGGLGRKEFWQLLSEGRATRGIT
KSVNZ -----MTARRVVTGIEVLAPGGTGSKAFWALLSEGRATRGIT
KSNOG -----MKESINRRVVTGIGIVAPDATGVKPFWOLLTAGRTATRTIT
KSTCM -----MTRHAEKRVVTGIGVAPGGAGTAAFWOLLTAGRTATRTIS
KSDAU -----MNRRVVTGMGVVAPGAIGIKSFWELLSGTTATRAIT
KSPEU -----MNRRVITGIGVAPGAVGTPKPFWELLSGTTATRAIS
KSWHI -----MTRRVAVVTGIGVAPGGIGTPQFWRLLESGRTATRRIS

```

\*\*\*: : : \* . \*

```

KCLFDAU RFADDGRLGRLAGEVSDFP-EDHLPKRLLVQTDPMQMTALAAEWALREAGCAPSS--
KCLFPEU RFTGDGRLGRLAGEVSDFP-EDHLPKRLLAQTDPMQY-ALAAAEWALRESGCCSPSS--
KCLFACT RFDVSRYPATLAGQIDDFHA-PDHPGRLLPQTD PSTRL-ALTAADWALQDAKADPES-L
KCLFHIR RFDPTGYPAELAGQVLDFA-TEHLPKRLLPQTDVSTRF-ALAAAALADAEVDPAE-L
KCLFGRA RFDASRYPSRLAGQIDDFEA-SEHLPKRLLPQTDVSTRY-ALAAADWALADAGVGPESGL
KCLFNOG RFDAGRYPSKLAGVPGFVP-EDHLPKRLLPQTDHMTL-ALVAADWAFQDAAVDP SK-L
KCLFTCM RFDPHGYPVVVGGEVLAFDA-AAHLPGRLLPQTD RMTQH-ALVAAEWALADAGLEPEK-Q
KCLFCIN RFDPSGYPAQLAGEIPGFRA-AEHLPGRLVPQTD RVT RL-SLAAADWALADAGVEVAA-F
KCLFVNZ RFDPTGYPARLAGEVPGFAA-EEHLPKRLLPQTD RMTL-ALVAADWALADAGVRPEE-Q
KCLFWHIE REGCAHLPLRVAGEVHGFA-AETVEDRFLVQTD RFTHF-ALSATQH ALADARFGRADVD
KSGRA FFDASPFERSRIAGEI-DFDAVAGFGSPREVRMDRATQF-AVACTRDALADSGLD TGA-L
KSHIR FFDPTPNRSQIAAEC-DFDPEHEGLSPREIRMDRAAQF-AVVCTRDADVADSGLEFEQ-V
KSACT FFDPSFYRSQVAAEA-DFDPVAGFGPRELDRMDRASQF-AVACAREFAASGLDPDT-L
KSCIN FFDPAFPERSKVAAEA-DFCGLENGLSPOEVRMDRAAQF-AVV TAR-AVEDSGAELAA-H
KSVNZ FFDPTPFRSRVAAEI-DFDPEAHGLSPQEIRMDRAAQF-AVVAAR-AVADSGIDLAA-H
KSNOG AFDPSPFRSRIAAEC-DFDPLAEGLT PQQIRMDRATQF-AVVSARESLED SGLDLGA-L
KSTCM LFDAAPYRSRIAGEI-DFDPIGEGLSPQASTYDRATQL-AVVCAREALKDSGLDPAA-V
KSDAU TFDATPFRSRIAAEC-DFDPVAAGLSAEQARRLD RAGQF-ALVAGQEAL TDSGLRIGE-D
KSPEU TFDATPFRSRIAAEC-DFDPVAAGLSAEQARRLD RAGQF-ALVAGQEALADSGLRIDE-D
KSWHI LFDPSGLRSQIAAEC-DFEPSDHGLGLATAQRCDRYVQF-ALVAASEAVRDANLDMNR-E

```

...: \* . \* : : : . :

Fig 2A

KCLFDAU  
KCLFPEU  
KCLFACT  
KCLFHIR  
KCLFGRA  
KCLFNOG  
KCLFTCM  
KCLFCIN  
KCLFVNZ  
KCLFWHIE  
KSGRA  
KSHIR  
KSACT  
KSCIN  
KSVNZ  
KSNOG  
KSTCM  
KSDAU  
KSPEU  
KSWHI

-PLEAGVITASASGGFASGQRELQNLWSKG-----PAHVSAYMSFAWFY-AVNTGQIAIR  
-PLEAGVITASASGGFAFGQRELQNLWSKG-----PAHVSAYMSFAWFY-AVNTGQIAIR  
TDYDMGVVTANACGGFDFTHREFRKLWSEG-----PKSVSVYESFAWFY-AVNTGQISIR  
PEYGTGVIITSNATGGFEFTHREFRKLWAQG-----PEFVSVYESFAWFY-AVNTGQISIR  
DDYDLGVVTSTAQGGFDFTHREFHKLWSQG-----PAYVSVYESFAWFY-AVNTGQISIR  
PEYGVGVVTASSAGGFEGFHRELQNLWSLG-----PQYVSAYQSFADFY-AVNTGQVSIR  
DEYGLGVLTAAGAGGFEGQREMYRQKLWGTG-----PERVSAYQSFADFY-AVNTGQISIR  
DPLDMGVVTASHAGGFEGQDELQKLLGQG-----QPVL SAYQSFADFY-AVNSGQISIR  
DDFTMGVVTASASGGFEGQDELQKLSQG-----SQYVSAYQSFADFY-AVNSGQISIR  
SPYSVGVTAAAGSGGGFEGQRELQNLWGHG-----SRHVGPYQSIADFY-AASTGQVSIR  
DPSRIGVALGSASASATSLENEYLVMDSGSEWLVDPALHSPMFDYLSPGVMPAEVAVA  
PPERIGVSLGSASAAATSLERQYLLVSDGGREWQVDPAYLSAHMFDYLSPGVMPAEVAWT  
DPAVGVSLGSASAAATSLEREYLLVSDSGRDWEVDAWLSRHMFDYLVPSVMPAEVAVA  
PPHRI GVVGSASVAGTNEYRVVSDGGRLDLVDHRYAVPHLYNYLVPSVMPAEVAVA  
DPYRVGVTVGSASVAGTMDLDEEYRVVSDGGRLDLVDHAYAVPHLYNYLVPSVMPAEVAVA  
DASRTGVVGSASVAGCTTSLLEEYAVVSDSGRNWLVDDGYAVPHLYNYLVPSVMPAEVAVH  
NPERIGVSI GTAVGCTTGLDREYARVSEGGSRWLVDDHILAVEQLEFDYFVPTSICREVAWE  
SAHRVGVCGTAVGCTQKLESEYVALSAGGANWVDPHGRGAPELYDYFVPSVMPAEVAVL  
SAHRVGVCGTAVGCTQKLESEYVALSAGGAHWVDPGRGSPELYDYFVPSVMPAEVAVL  
DPWRAGATLGTAVGGTTRLEHDYVLVSEGRWVDDRRSEPHLERAFATPATSASVAAEE

\* . . . . . : : : : \*



KCLFDAU  
KCLFPEU  
KCLFACT  
KCLFHIR  
KCLFGRA  
KCLFNOG  
KCLFTCM  
KCLFCIN  
KCLFVNZ  
KCLFWHIE  
KSGRA  
KSHIR  
KSACT  
KSCIN  
KSVNZ  
KSNOG  
KSTCM  
KSDAU  
KSPEU  
KSWHI

-HDLRGPVGVVVAEQAGGLDALAHAR-RKVRGGAE-LIVSGAMDSSSLCP-YGMAAQVRSR  
-HDLRGPVGVVVAEQAGGLDALAHAR-RKVRGGAE-LIVSGAVDSSSLCP-YGMAAQVKS  
-HGMRGPSALVAEQAGGLDALGHAR-RTIRRGTP-LVVSGGVDSALDP-WGWVSQIASG  
-HGLRGPVSVLVAEQAGGLDAVGHGG--AVRNGTP-MVVTGGVDSSFDL-WGWVSHVSSG  
-NTMRGPSAALVGEQAGGLDAIGHAR-RTVRRGPG-WCSAVASTRRSTR-GASSSQLSGG  
-HGLRGPVGVVLTVEQAGGLDALQAR-RQLRRGLP-MVAVAGVDGSPCP-WGWVAQLSSG  
-HGMRGHSSVFTVEQAGGLDAAHAA-RLLRKGTLNTALTGGCEASLCP-WGLVAQIPSG  
-HGMKGPSGVVSDQAGGLDALAQAR-RLVRKGTP-LIVCGAVEPRAPGAGSPSSPAGG  
-NGMKGPSGVVSDQAGGLDAVAQAR-RQIRKGTR-LIVSGGVDSALCP-WGWVAHVASD  
-NDFKGPVGVVVAEAGGLDALAHAA-LAVRNGTD-TVVCGATEAPLAP-YSIVCQLGYP  
-AGAEGPVTVSDGCTSGLDVSVGYAV-QGTREGSADVVAAGADTPVSPVIVACFDAIKA  
-VGAEGPVAMVSDGCTSGLDVSVGNV-RAIEEGSADVMFAGAADTPITPIVACFDAIRA  
-VGAEGPSTVSTGCTSGLDVSVGNV-RAIEEGSADVMFAGAADTPITPIVACFDAIRA  
-VGAEGPNTVSTGCTSGLDVSVGYARGELIREGSADVMFAGAADTPITPIVACFDAIRA  
RIGAEGPVSLVSTGCTSGLDVAVGRAA-DLIAEGAADVMLAGATEAPISPIVACFDAIKA  
-AGAEGPVTVSTGCTSGLDVAVGYGT-ELIRDGRADVVCATDAPISPIVACFDAIRA  
-AGAEGPVNI VSAGCTSGIDSIGYAC-ELIREGTVDVMLAGGVDAPIAPITVACFDAIRA  
-AGAEGPVNI VSAGCTSGIDSIGYAC-ELIREGTVDVMLAGGVDAPIAPITVACFDAIRA  
-FGVRGPVQTVSTGCTSGLDVAVGYAY-HAVABGRVDVCLAGAADSPISPIVACFDAIKA

\* . . . . . : : : : \*



KCLFDAU  
KCLFPEU  
KCLFACT  
KCLFHIR  
KCLFGRA  
KCLFNOG  
KCLFTCM  
KCLFCIN  
KCLFVNZ  
KCLFWHIE

RLSGSDOPTAGYLPFDRRAAGHVPGEG-GAILAVEDAERVAERG-GKVYGSIAGT-ASFD  
RLSGSDNPTAGYLPFDRRAAGHVPGEG-GAILTVEDAERAAERG-AKVYGSIAGYGASFD  
RI STATDPDRAYLPFDERAAGYVPGEG-GAILVLEDSAAAEARGRHDA YGELAGCASTFD  
RVSRATDPGRAYLPFDVAANGYVPGEG-GAILLLEDAESAKARG-ATGYGEIAGYAATFD  
LVSTVADPERAYLPFDVDASGYVPGEG-GAVLIVEDADSARARG---AERYVRSPLRRD  
GLSTSDPPRAYLPFDAAAGGHVPGEG-GALLVLESDESARARGVTRWYGRIDGYAATFD  
FLSEATDPHDA YLPFDARAAGYVPGEG-GAMLVAERADSARERDAATVYGRLAGHASTFD  
-MSDSDEPNRAYLPFDRDGRGYVPCGGRGVPPLERAEAAPARG-AEVYGE-AGPLARL-  
RLSTSEEPARGYLPFDREAQGHVPGEG-GAILVMEAAEAARERG-ARIYGEIAGYGSTFD  
ELSRATEPDRAYRPFTEAACGFAPAEG-GAVLVVEEEAAARERG-ADVRATVAGHAATFT

Fig 2B

KSGRA  
KSHIR  
KSACT  
KSCIN  
KSVNZ  
KSNOG  
KSTCM  
KSDAU  
KSPEU  
KSWHI

TTPRNDPAHASRPFDRNGFVLAEG-AAMFVLEEEYAAQRRG-AHIYAEVGGYATRSQ  
TTPRNDPEHASRPFDRNGFVLAEG-AALFVLEEEHARARG-AHVYAEISGCATRLN  
TTARNDDPEHASRPFDRNGFVLAEG-AAMFVLEEDYDSALARG-ARIHAEISGYATRCN  
TTPRHDAPATASRPFDRNGFVLGEG-AAFFVLEELHSARRRG-AHIYAEIAGYATRSN  
TTNRYDDPAHASRPFDRNGFVLGEG-AAVFLVEELHSARRRG-AHIYAEIAGYATRSN  
TTPRNDTPAEASRPFDRNGFVLGEG-AAVFLVEEFHARRRG-ALVYAEIAGFATRCN  
TSANNDPAHASRPFDRNGFVLGEG-SAVFVLEELSAARRRG-AHAYAEVRGFATRSN  
TSDHNDTPETLA-PFSRSRNGFVLGEG-GAIVVLEEEAAVRRG-ARIYAEIGGYASRGN  
TSDHNDTPETASRPFDRNGFVLGEG-GAIVVLEEEAAVRRG-ARIYAEIGGYASRGN  
TSPNNDPAHASRPFDRNGFVMGEG-AAVLVLEEDLEHARARG-ADVCEVSGYATFCN

\* \*\* \* . . \* . . \*

KCLFDAU  
KCLFPEU  
KCLFACT  
KCLFHIR  
KCLFGRA  
KCLFNOG  
KCLFTCM  
KCLFCIN  
KCLFVNZ  
KCLFWHIE  
KSGRA  
KSHIR  
KSACT  
KSCIN  
KSVNZ  
KSNOG  
KSTCM  
KSDAU  
KSPEU  
KSWHI

-PPPGSGRP---SALARAVETALADAGLDRSDIADVADGAA-VGELDVAEAEALASVFG  
-PPPGSGRP---SALARAVETALADAGLDSGLADVADGAA-VPELDAAEAEALASVFG  
-PAPGSGRP---AGLERAIRLALNDAGTGPEDVDVVFADGAG-VPELDAAEARAIGRVFG  
-PAPGSGRP---PALRRRAIELALADAELRPEQVDVVFADAG-VPELDAAEARAIGRVFG  
-PAPGSGRP---PALGRAAEALAEAGLTPADISVVFADGAG-VPELDRAEADTLARLFG  
-PPPGSGRP---PNLLRAAQALDDAEVGPEDVDVVFADAG-VPELDRAEADTLARLFG  
-ARPGTGRP---TGPARAIRLALDEARVAPEDVDVVFADAG-VPELDRAEADTLARLFG  
-PAPGSGRP---STRAHAIRLALDDAGTAPGDIRRVFADGGGRYPN-DRAEAEAEISEVFG  
-PRPGSGRE---PGLRKAIELALADAGAAPGDIADVADGAA-VPELDRAEAEALNAVFG  
GAGRWAESR---EGLARAIQALAEAGCRPEEVDVVFADAG-VPELDRAEAEALNAVFG  
-AYHMTGLKKGREMAESIRALDEARLDRTAVDYVNAHGSG-TQONDRHETAFAFKRSLG  
-AYHMTGLKTDGREMAEIRVALDLARIDPTDIDYINAHGSG-TQONDRHETAFAFKRSLG  
-AYHMTGLKADGREMAETIRVALDESRTDATDIDYINAHGSG-TQONDRHETAFAFKRSLG  
-AYHMTGLR-DGAEMAEAIRLALDEARLNPEQVDYINAHGSG-TQONDRHETAFAFKRSLG  
-AYHMTGLRPDGAEMAEIRVALDEARMNPTEIDYINAHGSG-TQONDRHETAFAFKRSLG  
-AFHMTGLRPDGAEMAEIATAALDQARRTGDDLYINAHGSG-TQONDRHETAFAFKRSLG  
-AYHMTGLRADGAEMAAAIATAALDEARRDPADVYVNAHGTA-TQONDRHETAFAFKRSLG  
-AYHMTGLRADGAEMAAAIATAALDEARRDPADVYVNAHGTA-TQONDRHETAFAFKRSLG  
-AYHMTGLTKEGLEMARAIATAALDEARRDPADVYVNAHGSG-TQONDRHETAFAFKRSLG

: : \* : : : \* : : \*

Fig 2c

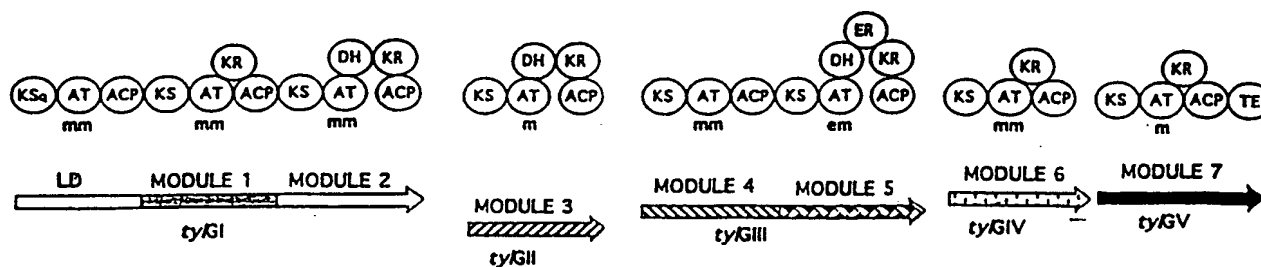
KCLFDAU P--HRVPVTVPKTLTGRLYSGAGPLDVATGLLALRDEVVPATGHVH-PDPDLPLDVVTGR  
 KCLFPEU P--RRVPVTVPKTLTGRLYSGAGPLDVATALLALRDEVVPATAHVD-PDPDLPLDVVTGR  
 KCLFACT R--EGVPVTVPKTTTGRLYSGGGPLDVVTALMSLREGVIAPTAGVTSVPREYGIDLVLGE  
 KCLFHIR P--SGVPVTAPKIMTGRLYSGGGPLDLVAALLAIRDQVI PPTVHTAEPVPEHQDLVTGD  
 KCLFGRA P--RGVPVTAPKALTGRLCAGGGPADLAAALLALRDQVI PATGRHRAVPDAYALDLVTGR  
 KCLFNOG P--YGVVTVAPKIMTGRLSAGGAALDVATALLALREGVVPPTVNVSRPRPEYELDLVLA-  
 KCLFTCM P--GAVPVTAPKIMTGRLYAGGAALDVATALLSIRDVVPPTVGTGAPAGLGIDLVLHQ  
 KCLFCIN P--GRVPVTCPRIMTGRLLHSGAAPLDVACALLAMRAGVI PPTVHID-PCPEYDLDLVLYQ  
 KCLFVNZ T--GAVPVTAPKIMTGRLYSGAAPLDLAAAFLAMDEGVI PPTVNVE-PDAAYGLDLVVGG  
 KCLFWHIE PHAARVPVTAPKTGTGRAYCAAPVLDVATAVLAMEHGLIPPTPHVL--DVCHDLDLVTGR  
 KSGRA EHAYAVPVSSIKSMGGHSLGAIGSIEIAASVLAIEHNVVPPTANLHTPDPECDLDYVPLT  
 KSHIR EHAYRTPVSSIKSMVGHSLGAIGSIEVAACALAI EHGVVPPTANLHEPDPECDLDYVPLT  
 KSACT EHARRTPVSSIKSMVGHSLGAIGSLEIAACVLALEHGVVPPTANLRTSDPECDLDYVPLE  
 KSCIN EHAYRTPVSSIKSMVGHSLGAIGSIEIAASALAMEYDVVPPTANLHTPDPECDLDYVPLT  
 KSVNZ DHAYRTPVSSIKSMVGHSLGAIGSIEIAASVLAITHDVVPPTANLHEPDPECDLDYVPLR  
 KSNOG DHAYRVPVSSIKSMIGHSLGAIGSLEIAASVLAITHDVVPPTANLHEPDPECDLDYVPLR  
 KSTCM QRAYDVPVSSIKSMIGHSLGAIGSLELAACALAI EHGVIPPTANYEEDPECDLDYVPLT  
 KSDAU DHAYRVPISSVKSMIGHSLGAAGSLEVAATALAVEYGAIPPTANLHDPDPELDLDYVPLT  
 KSPEU EHAYRVPISSVKSMIGHSLGAVGSLEVAATALAVEYGVIPPTANLHDPDPELDLDYVPLT  
 KSWHI EHAYATPMSSIKSMVGHSLGAIGSIEIAACVLAMAHQVVPPTANYTTPDPECDLDYVPRE  
 .\*:: :: \* : . ::: ::: :...\* : \* \*

KCLFDAU PRAMADARAALVVARGHGGFNSALVVRGAA-----  
 KCLFPEU PRSLADARAALLVARGYGGFNSALVVRGAA-----  
 KCLFACT PRSTAPRTA-LVLARGRWGFNSAAVLRRAFPTP----  
 KCLFHIR PRHQQLGTA-LVLARGKWGFNSAVVVRGVGTG-----  
 KCLFGRA PREAALSAA-LVLARGRHGFNSAVVVTLRGSDHRRPT  
 KCLFNOG PRRTPLARA-LVLARGRGGFNAAMVVAGPRAETR---  
 KCLFTCM PRELRVDTA-LVVARGMGGFNSALVVRRHG-----  
 KCLFCIN VRPAALRTA-LGGARGHGGFNSALVVRAGQ-----  
 KCLFVNZ PRTAEVNTA-LVLIARGHGGFNSAMVVRSAN-----  
 KCLFWHIE ARPAEPRTA-LVLARGLMGSNSALVLRRAVPPPEGR-  
 KSGRA AREQRVDTV-LTVSGSGFGGFQSAMVLRPEEAA----  
 KSHIR AREQRVDTV-LTVSGSGFGGFQSAMVLRRLGGANS---  
 KSACT ARERKLRVS-LTVSGSGFGGFQSAMVLRDAETAGAAA-  
 KSCIN ARDQRVDSV-LTVSGSGFGGFQSAMVLTSAQ---RSTV  
 KSVNZ CREQLTDSV-LTVSGSGFGGFQSAMVLRARPE---RKIA  
 KSNOG ARACPVDTV-LTVSGSGFGGFQSAMVLCGPGSRGRSAA  
 KSTCM AREQRVDTV-LTVSGSGFGGFQSAAVLARPKETRS---  
 KSDAU AREKVRHA-LTVSGSGFGGFQSAMLLSRPER-----  
 KSPEU AREKVRHA-LTVSGSGFGGFQSAMLLSRLER-----  
 KSWHI ARERTLRHV-LTVSGSGFGGFQSAVVLSGSEGGLR---  
 \* . \* . \* \* : \* :

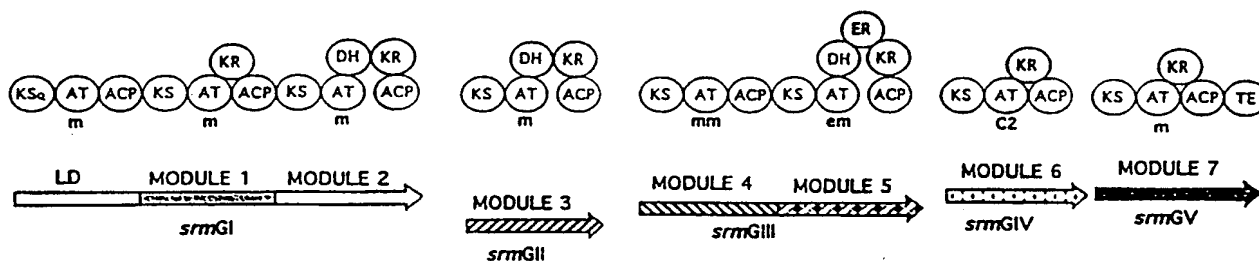
mole:-/ks2%

Fig 2D

## ORGANISATION OF THE TYLOSIN-PRODUCING POLYKETIDE SYNTHASE



## ORGANISATION OF THE SPIRAMYCIN-PRODUCING POLYKETIDE SYNTHASE



## ORGANISATION OF THE NIDDAMYCIN-PRODUCING POLYKETIDE SYNTHASE

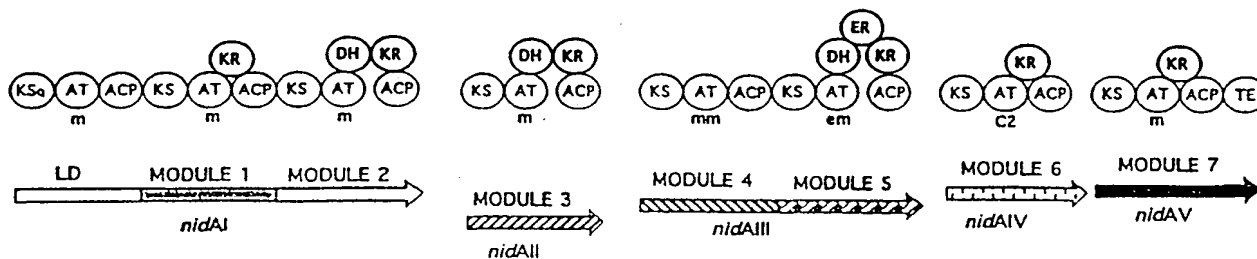


Fig 3

m: malonyl transferase  
 mm: methylmalonyl transferase  
 em: ethylmalonyl transferase  
 C2: unknown C2 unit transferase



Fig. 4A

	1					50
niddamycin	-----	-----	MAGHGDATAQ	KAQDAEKSED	GSDAIAVIGM	
platenolide	-----	-----	-----MS	GELAISRSDD	RSDAVAVVGM	
monensin	-----	-----	-----MAAS	ASASPSGPSA	GPDPIAVVGM	
oleandomycin	-----	-----	-----	---MHVPGEE	NGHSIAIVGI	
tylosin	MSSALRRAVQ	SNCGYGDLMT	SNTAAQNTGD	QEDVDGPDST	HGGEIAVVGM	
	51					100
niddam...	SCRFPGAPGT	AEFWQLLSSG	ADAVVTAADG	RRR.....	.....GTIDA	
platenol.	ACRFPGAPGI	AEFWKLLTDG	RDAIGRDADG	RRR.....	.....GMIEA	
monensin	ACRLPGAPDP	DAFWRLLESG	RSVSTAPPE	RRRADSGLHG	P...GGYLDR	
oleandom	ACRLPGSATP	QEFWRLLADS	ADALDEPPAG	RFPTGSLSSP	PAPRGGFLDS	
tylosin	SCRPGAAGV	EEFWELLRSG	RGMPTRQDDG	TWRAA.....	.....LED	
	101					150
niddam...	PADFDAFFG	MSPREAAATD	PQORLVLELG	WEALEDAGIV	PESLRGEAAS	
platenol.	PGDFDAFFG	MSPREAAETD	PQORLMLELG	WEALEDAGIV	PGSLRGEAVG	
monensin	IDGFDADFFH	ISPRAVAMD	PQORLLELS	WEALEDAGIR	PPTLARSRTG	
oleandom	IDTFDADFFN	ISPRAEAGVD	PQORLLELG	WEALEDAGIV	PRHLRGTRTS	
tylosin	HAGFDAGFFG	MNARQAAATD	PQHRLMLELG	WEALEDAGIV	PGDLTGTDGT	
	151					200
niddam...	VFVGAMNDDY	ATLLH.RAGA	PTDITYTATGL	QHSMIANRLS	YFLGLRGPSL	
platenol.	VFVGAMHDDY	ATLLH.RAGA	PVGPHATATGL	QRAMLANRLS	YVLGTRGPSL	
monensin	VFVGAFWDDY	TDVLNLRAPG	AVTRHTMTGV	HRSILANRIS	YAYHLAGPSL	
oleandom	VFMGAMWDDY	AHLAHARGEA	ALTRHSLTGT	HRGMIANRLS	YALGLQGPSL	
tylosin	VFAGVASDDY	A.VLTRRSV	SAGGYTATGL	HRLAANRLS	HFLGLRGPSL	
	201					250
niddam...	VVDTGQSSSL	VAVALAVESL	RGGTSGIALA	GGVNLVLAEE	GS.AAMERVG	
platenol.	AVDTAQSSSL	VAVALAVESL	RAGTSRVAVA	GGVNLVLADE	GT.AAMERLG	
monensin	TVDTAQSSSL	VAVHLACESI	RSGLSDIAFA	GGVNLICSPR	TTELAAARFG	
oleandom	TVDTGQSSSL	AAVHMACESE	ARGESDLALV	GGVNLVLDPA	GT.TGVERFG	
tylosin	VVDSAQSASL	VAVQLACESL	RRGETSLAVA	GGVNLILTEE	ST.TVMERMG	
	251					300
niddam...	ALSPDGRCHT	FDARANGYVR	GEGGAIVVLK	PLADALADGD	RVYCVVRGVA	
platenol.	ALSPDGRCHT	FDARANGYVR	GEGGAIVVLK	PLADALADGD	PVYCVVRGVA	
monensin	GLSAAGRCHT	FDARADGFVR	GEGGGLVVLK	PLAAARRDGD	TVYCVIRGSA	
oleandom	ALSPDGRCHT	FDSRANGYAR	GEGGVVVVLK	PTHRALADGD	TVYCEILGSA	
tylosin	ALSPDGRCHT	FDARANGYVR	GEGGGAIVVLK	PLDAALADGD	RVYCVIKGGA	
	301					350
niddam...	TGNDGGGPGL	TVPDRAGQEA	VLRAACDQAG	VRPADVRFVE	LHGTGTPAGD	
platenol.	VGNDGGGPGL	TAPDREGQEA	VLRAACAQAR	VDPAEVRFVE	LHGTGTPVGD	
monensin	VNSDGTDDGI	TLPSGQAQGD	VVRLACRRAR	ITPDQVQYVE	LHGTGTPVGD	
oleandom	LNNDGATEGL	TVPSARAQAD	VLRQAWERAR	VAPTDVQYVE	LHGTGTPAGD	
tylosin	VNNDGGGASL	TTPDREAQEA	VLRQAYRRAG	VSTGAVRYVE	LHGTGTRAGD	

	351		400
niddam...	PVEAEALGAV	YGTGRP..AN	EPLLVGSVKT NIGHLEGAAG IAGFVKAALC
platenol.	PVEAHALGAV	HGSGRP..AD	DPLLVGSVKT NIGHLEGAAG IAGLVKAALC
monensin	PIEAAALGAA	LGQDAA..RA	VPLAVGSAKT NVGHLEAAAG IVGLLKTALS
oleandom	PVEAEGLGTA	LGTARP..AE	APLLVGSVKT NIGHLEGAAG IAGLLKTVLS
tylosin	PVEAAALGAV	LGAGADSGRS	TPLAVGSVKT NVGHLEGAAG IVGLIKATLC
	401		450
niddam...	LHERALPASL	NFETPNPAIP	LERLRLKVQT AHAALQPGTG GGPLLAVGSA
platenol.	LRERTLPGSL	NFATPSPAIP	LDQLRLKVQT AAAELPLAPG GAPLLAGVSS
monensin	IHHRRLAPSL	NFTTPNPAIP	LADLGLTVQQ DLADWP..RP EQPLIAGVSS
oleandom	IKNRHLPASL	NFTSPNPRID	LDALRLRVHT AYGPWP..SP DRPLVAGVSS
tylosin	VRKGELVPSL	NFSTPNPDIP	LDDLRLRVQT ERQEW.NEED DRPRVAGVSS
	451		500
niddam...	FGMGGTNCHV	VLEETPGG..	.....RQPAE.T
platenol.	FGIGGTNCHV	VLEHLPSR..	.....PTPAV.S
monensin	FGMGGTNCHV	VVA....AAP	DSVAVPEPVG VPERVEVPEP VVVSEPVVVP
oleandom	FGMGGTNCHV	VLSELRNAGG	DGAGKGPYTG TEDRLGATEA EKRPDPATGN
tylosin	FGMGGTNVHL	VIAEAPAAAG	SSGAGGSGAG SGAGISAVSG VV.....
	501		550
niddam...	GQADACLFSA	SPMLLLSARS	EQALRAQAAR LREHL..EDS GADPLDIAYS
platenol.	VAAS...LPD	VPPLLLSARS	EGALRAQAVR LGETV..ERV GADPRDVAYS
monensin	TPWP.....	.....VSAHS	ASALRAQAGR LRTHLAAHRP TPDAARVGHA
oleandom	GPDPAQDTHR	YPALILSARS	DAALRAQAER LRHHL.EHSP GQRLRDTAYS
tylosin	.....	..PVVVSGRS	RVVVREAAGR LAE..VVEAG GVGLADVAVT
	551		600
niddam...	LATTRTRFEH	RAAVPCGDPD	RLSSALAALA AGQTPRGVRI GS..TDADGR
platenol.	LASTRTLFEH	RAVVPCGGRG	ELVAALGGFA AGRVSGGVRS GR..A.VPGG
monensin	LATTRAPLAH	RAVLLGGDTA	ELLGSLDALA EGAETASIVR GEAYT..EGR
oleandom	LATRRQVFER	HAVVTGHDRE	DLLNGLRDLE NGLPAPQVLL GRTPTPEPGG
tylosin	MAD.RSRFGY	RAVVLARGEA	ELAGRLRALA GGDPDAGVVT G...AVLDGG
	601		650
niddam...	LALLFTGQGA	QHPGMQOELY	TTDPHFAAAL DEVCEELQRC GTQNLREVMF
platenol.	VGVLFTGQGA	QWVGMGRGLY	AGGGVFAEVL DEVLSMVGEV DGRSLRDVMF
monensin	TAFLFSGQGA	QRLGMGRELY	AVFPVFADAL DEAFALDVH LDRPLREIVL
oleandom	LAFLFSGQGS	QQPGMGKRLH	QVFPGFRDAL DEVCAELDTH LGRLL.....
tylosin	VVVGAAAPGA	GAAGGAGAAG	GAGGGGVVLV FPGQGTQWVG MGAGLLGSSE
	651		700
niddam...	TPDQPD....	.....	LLDRTEYTQP ALFALQTALY
platenol.	GDVDVDAGAG	ADAGAGAGAG	VGSGSGSVGG LLGRTEFAQP ALFALEVALF
monensin	GETDSGGNVS	GENVIGEGA..	.....DHQA LLDQTAYTQP ALFAIETSLY
oleandom	.GPEAGPPLR	DVMFAERGT.	.....AHSA LLSETHYTQA ALFALETALF
tylosin	VFAASMRECA	RALSVHVGWD	LLEVSGGAG .LERVDVVQP VTWAVMVSLA
	701		↓ 750
niddam...	RTLTARGETQA	HLVLGHSVGE	ITAHIAGVL DLPDAARLIT ARAHVMGQLP
platenol.	RALEARGVEV	SVVLGHSVGE	VAAATVAGVL SLGDAVRLVV ARGGLMGGLP
monensin	RLAASFGLKP	DYVLGHSVGE	IAAAHVAGVL SLPDASALVA TRGRMLQAVR
oleandom	RLLVQWGLKP	DHLAGHSVGE	IAAAHAAGIL DLSDAAEVLA TRGALMRSLP
tylosin	RYWQAMGVDV	AAVVGHSQGE	IAAATVAGAL SLEDAAVVA LRAGLIGRYL

Fig 4B

	751		800
niddam...	HG.GAMLSVQ	AAEHDLDQLA	HTHG..VEIA AVNGPTHCVL SGPRTALEET
platenol.	VG.GGMWSVG	ASESVVRGVV	EGLGEWVSVA AVNGPRSVVL SGDVGVLSEV
monensin	AP.GAMAAWQ	ATADEAAEQ	AGHERHVTVA AVNGPDSVVV SGDRATVDEL
oleandom	GG.GVMLSVQ	APESVAPLL	LGREAHVGLA AVNGPDVAVV SGERGHVAAI
tylosin	AGRGAMAAVP	LPAGEVEAGL	.AKWPGVEVA AVNGPASTVV SGDRRAVAGY
	801		850
niddam...	AQHLREQNVR	HTWLKVSHAF	HSALMDPMLG AFRDTLNTLN Y..QPPTIPL
platenol.	VASLMGDGVE	YRRLDVSHGF	HSVLMPEVLG EFRGVVESLE FGRVRPGVVV
monensin	TAAWRGRGRK	AHHLKVSHAF	HSPHMDPILD ELRAVAAGLT FHE..PVIPV
oleandom	EQILRDRGRK	SRYLRVSHAF	HSPLMEPVLE EFAEAVAGLT FRA..PTTPL
tylosin	VAVCQAEQVQ	ARLIPVDYAS	HSRHVEDLKG ELERVLSGI. .RPRSPRPV
	851		900
niddam...	ISNLTGQIA.	.....DPNHL	CTPDYWIDHA RHTVRFADAV QTAHHQGT
platenol.	VSGVSGGVV.	.....GSGEL	GDPGYWVRHA REAVRFADGV GVVRLGLVGT
monensin	VSNVTGELVT	ATATGSGAGQ	ADPEYWARHA REPVRFLSGV RGLCERGVT
oleandom	VSNLTG....	..APVDDRTM	ATPAYWVRHV REAVRFGDGI RALGKLGTGS
tylosin	CSTVAGEQPG	EPVF.....	.DAGYWFRNL RNRVEFSAVV GGLLEEHR
	901		950
niddam...	YLEIGPHPTL	TTLHHTL..	.DNP.....T TIPTLHRERP
platenol.	LVEVGPHGVL	TGMAGECLGA	GDDV.....V VVPAMRRGRA
monensin	FVELGPDAPL	SAMARDCFPA	P.....ADRSRPRPA AIATCRRGRD
oleandom	FLEVGPDGVL	TAMARACVTA	APEPGHRGEQ GADADAHTAL LLPALRRGRD
tylosin	FIEVSAHPVL	V.....	.....HAIEQ TAEAADRSVH ATGTLRRQDD
	951		
niddam...	EPETLTQAIA	AVGVRTDGID	WAVLCGASRP RRVELPTYAF
platenol.	EREVFEEAALA	TVFTRDAGLD	ATALHTGSTG RRIDLPTTFF
monensin	EVATFLRSLA	QAYVRGADVD	FTRAYGATAT RRFPLPTYPF
oleandom	EARSLTEAVA	RLHLHGVPMD	WTSVLGGDVS .RVPLPTYAF
tylosin	SPHRLLTSTA	EAWAHGATLT	WDPAL..PPG HLTTLPTYPF

niddam: niddamycin; platenol: platenolide I (spiramycin); oleandom: oleandomycin.

FIG. 4C

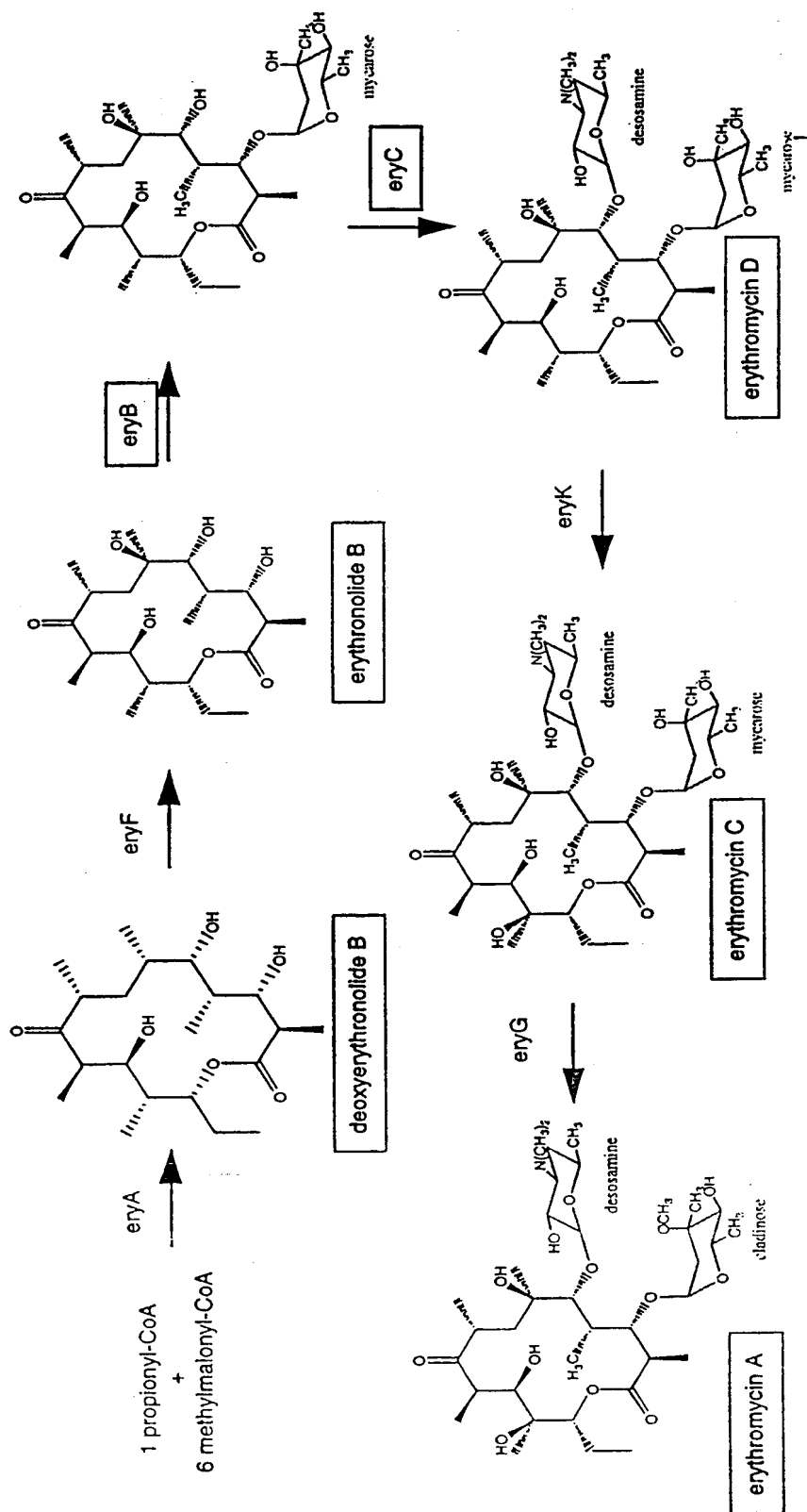


Fig. 5

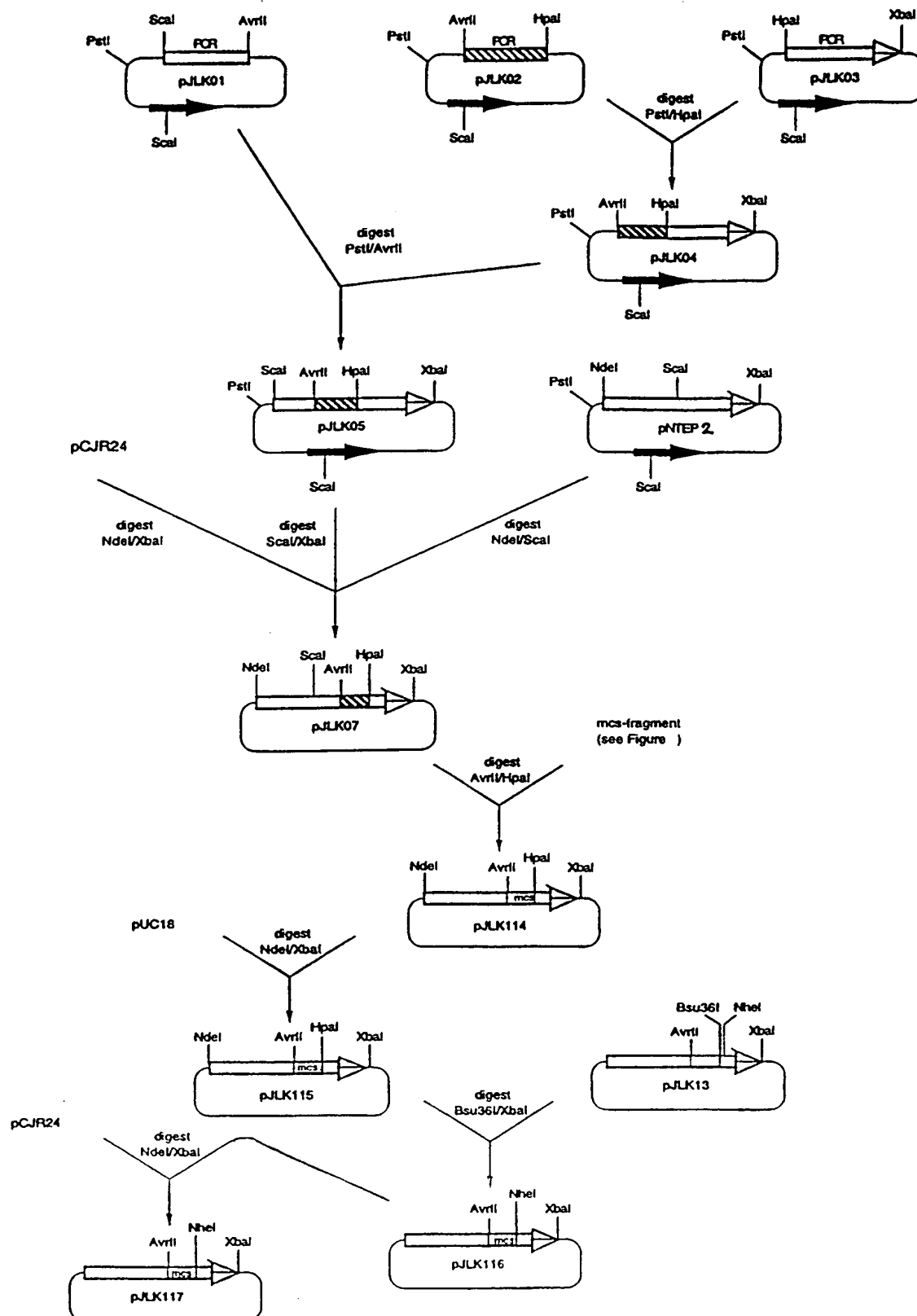


Fig 6

## Figure 7

forward (Plf):

5'-CTA GGC CGG GCC GGA CTG GTA GAT CTG CCT ACG TAT CCT TTC CAG GGC AAG CGG TTC TGG CTG CAG CCG GAC CGC ACT AGT CCT CGT GAC GAG

GGA GAT GCA TCG AGC CTG AGG GAC CGG TT-3'

backward (Plb):

5'-AAC CGG TCC CTC AGG CTC GAT GAT GCA TCT CCC TCG TCA CGA GGA CTA CTG CGG TCC GGC TGC AGC CAG AAC CGC TTG CCC TGG AAA GGA TAC GTA

GGC AGA TCT ACC AGT CCG GCC CGG C-3'

oligos annealed:

CTAGCCCGGCCGACTGGTAGATCTGCCTACGTATCCTTCCAGGGCAAGCGGTCTGGCTGCAGCCGACCGCCTAGTCTCGTGACGAGGGAGATGCATCGAGCCTGAGGGACCGGTT  
CGGCCCGGCCCTGACCATCTAGACCGATGCATAGGAAGGTCCCGTTGCGCAAGACCGACGTCGCGCTGGCGTGATCAGGACACTGCTCCCTACGTAGCTCGGACTCCCTGGCCAA

-----  
AvrII BglII SnaBI PstI SpeI NsiI Bsu36I HpaI  
-----



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>7</sup> :</b> <b>C07H 17/08, A61K 31/70</b>	<b>A3</b>	<b>(11) International Publication Number:</b> <b>WO 00/00500</b> <b>(43) International Publication Date:</b> 6 January 2000 (06.01.00)
<b>(21) International Application Number:</b> PCT/GB99/02042 <b>(22) International Filing Date:</b> 29 June 1999 (29.06.99)  <b>(30) Priority Data:</b> 9814006.4                      29 June 1998 (29.06.98)                      GB  <b>(71) Applicants (for all designated States except US):</b> BIOTICA TECHNOLOGY LIMITED [GB/GB]; 112 Hills Road, Cambridge CB2 1PH (GB). PFIZER INC. [US/US]; 235 East 42nd Street, New York, NY 10017 (US).  <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> LEADLAY, Peter, Francis [GB/GB]; 17 Clarendon Road, Cambridge CB2 2BH (GB). STAUNTON, James [GB/GB]; 29 Porson Road, Cambridge CB2 2ET (GB). CORTES, Jesus [GB/GB]; 26 Cambanks, Union Lane, Cambridge CB4 1PZ (GB). MCARTHUR, Hamish, Alastair, Irvine [GB/US]; 19 Pheasant Run Drive, Gales Ferry, CT 06335 (US).  <b>(74) Agents:</b> STUART, Ian et al.; Mewburn Ellis, York House, 23 Kingsway, London WC2B 6HP (GB).		<b>(81) Designated States:</b> AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>  <b>(88) Date of publication of the international search report:</b> 16 March 2000 (16.03.00)
<b>(54) Title:</b> POLYKETIDES AND THEIR SYNTHESIS  <b>(57) Abstract</b>  <p>A polyketide synthase ("PKS") of Type I is a complex multienzyme including a loading domain linked to a multiplicity of extension domains. This first extension module receives an acyl starter unit from the loading domain and each extension module adds a further ketide unit which may undergo processing (e.g. reduction). We have found that the Ksq domain possessed by some PKS's has decarboxylating activity, e.g. generating (substituted) acyl from (substituted) malonyl. The CLF domain of type II PKS's has similar activity. By inserting loading modules including such domains into PKS's not normally possessing them it is possible to control the starter units used.</p>		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MV	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						



# INTERNATIONAL SEARCH REPORT

Interr      1st Application No

PCT/GB 99/02042

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7    C07H17/08    A61K31/70

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7    C07H    A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>WO 91 16334 A (ABBOTT LAB)  31 October 1991 (1991-10-31)  the whole document, but especially page 7,  6-Deoxy-15-norerythromycins A-D. More  especially 6-Deoxy-15-norerythromycins A  and C</p> <p style="text-align: center;">--- -/--</p>	1-9

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

5 January 2000

Date of mailing of the international search report

20/01/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Scott, J

# INTERNATIONAL SEARCH REPORT

Intern al Application No

PCT/GB 99/02042

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	L.LIN ET AL.: "Biosynthesis of 2-Nor-6-deoxyerythronolide B by Rationally Designed Domain Substitution." JOURNAL OF THE AMERICAN CHEMICAL SOCIETY., vol. 119, no. 43, 29 October 1997 (1997-10-29), pages 10553-10554, XP002121612 AMERICAN CHEMICAL SOCIETY, WASHINGTON, DC., US ISSN: 0002-7863	1-14
X	the whole document, but especially compound 3	1,4,7
Y	WO 98 01546 A (CORTES JESUS ;LEADLAY PETER F (GB); STAUNTON JAMES (GB); BIOTICA T) 15 January 1998 (1998-01-15) cited in the application	1-17
X	the whole document	1,4,7
Y	WO 98 01571 A (CORTES JESUS ;LEADLAY PETER F (GB); PACEY MICHAEL S (GB); STAUNTON) 15 January 1998 (1998-01-15) cited in the application the whole document	1-17
Y	WO 95 08548 A (UNIV LELAND STANFORD JUNIOR ;JOHN INNES CENTRE (GB)) 30 March 1995 (1995-03-30) cited in the application the whole document	1-17
P,X	WO 98 49315 A (KOSAN BIOSCIENCES INC ;UNIV LELAND STANFORD JUNIOR (US)) 5 November 1998 (1998-11-05) the whole document, but especially claims 1-28 and figure 6/table 2, all compounds containing "nor"	1-17
E	WO 99 35157 A (WU YONG JIN ;PFIZER PROD INC (US)) 15 July 1999 (1999-07-15) the whole document	1,4,7

# INTERNATIONAL SEARCH REPORT

information on patent family members

Intern al Application No

PCT/GB 99/02042

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9116334 A	31-10-1991	US 5141926 A CA 2080583 A EP 0525083 A JP 2587562 B JP 5504890 T KR 9608668 B PT 97390 A	25-08-1992 19-10-1991 03-02-1993 05-03-1997 29-07-1993 28-06-1996 31-01-1992
WO 9801546 A	15-01-1998	AU 3450997 A AU 3451497 A CA 2259420 A CA 2259463 A CN 1229438 A EP 0909327 A EP 0910633 A WO 9801571 A GB 2331518 A NO 990012 A PL 331285 A AU 7666198 A WO 9854308 A	02-02-1998 02-02-1998 15-01-1998 15-01-1998 22-09-1999 21-04-1999 28-04-1999 15-01-1998 26-05-1999 23-02-1999 05-07-1999 30-12-1998 03-12-1998
WO 9801571 A	15-01-1998	AU 3450997 A AU 3451497 A CA 2259420 A CA 2259463 A CN 1229438 A EP 0909327 A EP 0910633 A WO 9801546 A GB 2331518 A NO 990012 A PL 331285 A AU 7666198 A WO 9854308 A	02-02-1998 02-02-1998 15-01-1998 15-01-1998 22-09-1999 21-04-1999 28-04-1999 15-01-1998 26-05-1999 23-02-1999 05-07-1999 30-12-1998 03-12-1998
WO 9508548 A	30-03-1995	US 5672491 A AU 678058 B AU 7731794 A CA 2171629 A EP 0725778 A JP 9505983 T US 5712146 A US 5830750 A US 5843718 A US 5962290 A	30-09-1997 15-05-1997 10-04-1995 30-03-1995 14-08-1996 17-06-1997 27-01-1998 03-11-1998 01-12-1998 05-10-1999
WO 9849315 A	05-11-1998	AU 7172298 A	24-11-1998
WO 9935157 A	15-07-1999	AU 1501599 A	26-07-1999

**THIS PAGE BLANK (USPTO)**